


```

STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA: /07/951.715A
TELEPHONE NUMBER: 919-511-8689
FILING DATE: 25-SEP-1992
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Spruill, W. Murry
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: S-18805/A/COC 1577/C1P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-511-8689
TELEFAX: 919-511-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYDROLYTIC DEGRADATION: /desc = "Synthetic DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3621
OTHER INFORMATION: /product= "Full-length, maize
OTHER INFORMATION: optimized cryiB"
OTHER INFORMATION: /note= "Disclosed in Figure 6."
US-07-951-715A-6

alignment_scores:
Quality: 75.50 Length: 147
Ratio: 1.198 Gaps: 7
Percent Similarity: 42.857 Percent Identity: 25.850

alignment_block:
US-09-327-7500-13 x US-07-951-715A-6 ..
Align seg 1/1 to US-07-951-715A-6 from: 1 to: 3624
5 HsldglnuasnlglduMetgldglnPrMeGlnasnlglygluGluas 21
2121 CACACACGAGCAGACGACCTCTCCAGC .. CATCCACGAGCAGACGAGC 217
21 PARpGLeuuglygylgylgylgylgylHsldglnPrMeGlnasnlglygluGluas 38
2168 ACGGCTGTGGGGGACGAGCAACATCACCATCCACGAGGAGCAGCAGGTG 2217
38 lyGlnala.....ArgArgLeuAlapto 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2218 TTCAGAGAACTACGTGACCTCCGGCAGCTTCACACAGGTGTCACCC 2267
46 AsnPha.....ArgTpal 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2268 CACCTACTCTTACACAGAAGATCGCGCAGGAGCAGCTGGAAGGCTTACACCC 2317
50 alleProAsnArgGlnIntlleAsnAspGlyMetgylgylAspIyAsp.... 65
2310 GCTACACAGCTCCGGGGCTACATCGAGGACACCGACAGCTGGAGATCTAC 2367
66 .....AspMet 67
2368 CTGATCCGCTACAGCGCAGCAGCAGACCTCGGAGCTGCTCGCAGCA 2417

```

[illegible]

```

alignment_scores:
  Quality: 75.50      Length: 147
  Ratio: 1.198      Gaps: 7
  Percent Similarity: 42.857      Percent Identity: 25.050

alignment_block:
  US-09-327-750D-13 x US-08-459-595A-6 ..

Align seq 1/1 to: US-08-459-595A-6 from: 1 to: 3524

          5  HsGlcInGluuGluMetGluGlnPrometGlnaGluGluGlna 21
          |||||:::||||:||||:||||:||||:||||:||||:
2121  CACCACGACGACAGACAGCAACTTCACAG...CATCCACGACGACGACGAC 2167

```


(UTC) UNIV COLUMBIA NEW YORK.

Sato T;
WPI; 2001-061707/07.

New p75-neurotrophin receptor-associated cell death executor (NADE) and the gene encoding NADE, useful for modulating the activity of p75NTR and for detecting neurodegenerative diseases -

Disclosure; Fig 1; 134pp; English.

The present invention relates to a purified polypeptide capable of binding neurotrophin receptor (p75-NTR). The invention is useful for binding and modulating the activity of p75NTR. The peptide mediates NGR-induced apoptosis, which plays an important role in neurodegenerative diseases. The peptide of the invention and p75NTR are useful for inhibiting NF-kappaB activation in a cell or a subject, for inducing caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase and fragment nuclear DNA in a cell by co-expression of (1) and p75-NTR.

Sequence 891 BP; 251 A; 182 C; 224 G; 234 T; 0 other;

localalignment_scores:
Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block: 13 x AAF23529 ...

Align seg 1/1 to: AAF23529 from: 1 to: 891

1 MetAlaAnlleHisGlnAsnGluMetGluGlnProMetGlnAs 17
|||||
312 ATGCGCAATATTCACCGAGAAAGAGAGATGGACGAGCTATGCAGAA 361
|||||
17 nGlyGluGluAspArgProteuGlyGlyGlyGluHisGlnProAlaG 34
|||||
362 TGGAGAGAGAGACCCCTTTGGAGGAGGTGAGGCGCCAGCGCTGCAG 411
|||||
34 TAsnAlaArgGluAlaArgGluAlaArgGluAlaArgGluAlaArgGluAla 50
|||||

```

PR 07-JUN-1999; 99US-0327750.
XX (UYCO ) UNIV COLUMBIA NEW YORK.
XX Sato T;
XX
XX
XX WPI; 2001-061707/07.
XX
XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
XX the gene encoding NADE, useful for modulating the activity of p75NTR
XX and for detecting neurodegenerative diseases.
XX
XX Claim 12; Fig 1; 134pp; English.
XX
XX The present invention relates to a purified polypeptide capable of
XX binding and activating the p75NTR. The invention is useful for
XX binding and activating the p75NTR. The peptide mediates
XX NGF-induced apoptosis which plays a role in neurodegenerative
XX diseases. The peptide of the invention and p75NTR are useful for
XX inhibiting NF-kappaB activation in a cell or a subject for inducing
XX caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
XX and fragment nuclear DNA in a cell by co-expression of (1) and p75-NTR.
XX
XX Sequence 700 BP; 177 A; 188 C; 203 G; 132 T; 0 other;
XX
alignment_scores:
XX Quality: 564.50 Length: 124
XX Ratio: 5.086 Gaps: 0
XX Percent Similarity: 89.516 Percent Identity: 82.258
XX
alignment_block:
XX US-09-327-750D-13 x AAF23528 ..
XX
XX Align seg 1/1 to: AAF23528 from: 1 to: 700
XX
XX 1 MetAlaAsnIleHisGluGluGluMetGluGluProMetGlnAs 17
XX 177 ATGGCCAAATATTCACCGAGAAACGAGAGATGGAGCAGCCCTGCGAGAA 226
XX 17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGluProAlaG 34
XX 227 TGGAGCAGGAGACGCCCTGTGGGAGAGGTGGGCCACACCTCTGCTG 276
XX 34 IyAsn.....ArgArg 37
XX 277 CAACACACACACACACACACACACACACACACACACACACACACAC 326
XX
XX 38 GlyGluAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 54
XX 327 GGGCAGGCTGGCGGACTTCCCTTAACTTCGATGGGCCATTCACACAG 376
XX 54 gClnIleAspAspGlyMetGlyGlyAspGlyAspMetGluIlePheW 71
XX 377 CGAGATGATGACGGTGGGTGGAGATGGAGATGATATGGAAATGTTCA 426
XX
XX 71 etGluGluMetArgGluIleArgArgLysLeuArgGluGluGluLeuArg 87
XX 427 TGGAGGAGATGAGAGAGATCCGAGGAAGCTTAGGAGCTACAGCTGAGA 476
XX
XX 88 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 104
XX 477 AATGTCTACGCTCTTATGGGGAGCTGTCTAACCCACACACATCACCA 526
XX
XX 104 sAspGluPheCysLeuMetPro 111
XX 527 TGATCAATTCCTCTATGCT 548
XX
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT: AAC01005
seq_documentation_block:
XX ID AAC01005 standard; cDNA; 532 BP.
XX

```

```

AC AAC01005;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 1003.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX P-PSDB; AAG00999.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 1003; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified in the
XX sequence. The 5' ESTs were prepared from total human RNAs for cDNA
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors.
XX
XX Sequence 532 BP; 151 A; 121 C; 163 G; 91 T; 6 other;
XX
alignment_scores:
XX Quality: 537.00 Length: 104
XX Ratio: 5.424 Gaps: 0
XX Percent Similarity: 95.192 Percent Identity: 95.192
XX
alignment_block:
XX US-09-327-750D-13 x AAC01005 ..
XX
XX Align seg 1/1 to: AAC01005 from: 1 to: 532
XX
XX 1 MetAlaAsnIleHisGluGluGluMetGluGluProMetGlnAs 17
XX 219 ATGGCAATATTCACCGAGAAACGAGAGATGGAGCAGCTATGCGAA 268
XX
XX 17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGluProAlaG 34
XX 269 TGGAGGAGAGACGCCCTTGGGAGAGGTGGAGCGTACGCCCTGCGAG 318
XX
XX 34 IyAsnArgArgGlyGluAlaArgArgLeuAlaProAsnPheArgTrpAla 50
XX 319 GAAATCCAGCGGGAGAGAGCTCGCCGAVTTGCCCTTAATTTTCATGGGCC 368
XX
XX 51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspMe 67
XX

```



```
PR 29-JAN-1999; 99US-0118078.
XX (SEAR) SEARLE & CO G D.
XX Bunch RT, Curtis SM, Rodi CP, Morris DL;
XX WPI; 2000-505977/45.
XX New nucleic acid encoding a carcinogenic biomarker, induced by
XX phenobarbital treatment of rat hepatocytes, useful for identifying
XX carcinogenic compounds.
XX Claim 1; Page 73; 240pp; English.
XX AA87080 to AA87656 represent nucleic acid sequences (N1) encoding a
XX carcinogenesis biomarkers. The carcinogenesis biomarkers are induced by
XX treating rat hepatocytes with phenobarbital. The nucleic acids are
XX useful for identifying carcinogenic compounds. The nucleic acid molecules
XX can be used to derive probes and/or primers for detecting or inducing
XX carcinogenesis, respectively.
XX Sequence 187 BP; 39 A; 48 C; 40 G; 60 T; 0 other;
SQ
alignment_scores:
  Quality: 193.00      Length: 35
  Ratio: 5.514        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
  US-09-327-750D-13 x AA87147
  Align seg 1/1 to: AA87147 from: 1 to: 187
  77 TleArGtGlyLeuAGtGluLeuGlnLeuArgAnCysLeuArgIlele 93
  |||||
  2 ATCCGGAGAAAGCTTAGGAGCTGCAGTTGAGAAATTTGCTGCTATCT 51
  |||||
  93 UMeGtGlyLeuSerAnHsHisAspHisAspGluPheCysLeuM 110
  |||||
  52 TATGGGAGAGCTCTCTATTCACACACACATCNCATCAGTATTCCTTA 101
  |||||
  110 etPro 111
  |||||
  102 TCCT 106
seq_name: /STD2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143
seq_documentation_block:
  ID AAH45143 standard; cDNA; 792 BP.
  AC AAH45143;
  DT 07-SEP-2001 (first entry)
  DE Human brain expressed X-linked protein, hbex, coding sequence.
  XX Human; brain expressed X-linked protein; cytostatic; auditory; nontropic;
  KW hbex; dysmrioplasia; hereditary disease; cancer; tumor; deafness;
  XX X-chromosome-binding mental retardation; lissencephalous disease; ss.
  OS Homo sapiens.
  XX WO200140286-A1.
  PN 07-JUN-2001.
  PD 27-NOV-2000; 2000WO-CN00502.
  PF 30-NOV-1999; 99CN-0124179.
  XX (BIOR-) BIORAD GENE DEV LTD SHANGHAI.
  PA
  KW
```

PI Mao Y, Xie Y;
XX WPI: 2001-397944/42.
XX P-PSDB; AAB59224.
XX Isolated human brain-expressed X-linked polypeptide used to diagnose
XX and treat of dysmrioplasia, hereditary diseases, cancer, tumor,
XX deafness and X-chromosome-binding mental retardation
XX Claim 5; Page 22; 30pp; Chinese.
XX The present sequence is the coding sequence for a human brain-expressed
XX X-linked protein (hbex). hbex and its coding sequence are useful in the
XX diagnosis and treatment of dysmrioplasia, hereditary diseases, cancer,
XX tumours, deafness, X-chromosome-binding mental retardation and
XX lissencephalous disease. hbex is also useful for screening mimics,
XX agonists, or inhibitors, and in peptide fingerprinting identification.
XX hbex coding sequence can be used as primers or probes, or in producing
XX gene chips or microarrays.
XX Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;
SQ

alignment_scores:
 Quality: 187.50 Length: 113
 Ratio: 2.679 Gaps: 3
 Percent Similarity: 61.947 Percent Identity: 42.478
alignment_block:
 US-09-327-750D-13 x AAH45143
 Align seg 1/1 to: AAH45143 from: 1 to: 792
 1 MetAlaLeuIleHisGlnLeuAGtGluLeuMetGluGlnProMetGlnAs 17
 |||||
 212 ATGGAAATGCCACCAAGAAATGAGAAAGGAGCAAGTTGCTAATAA 261
 |||||
 17 nGlyGluLu...AspArgProLeuGlyGlyGlyGlyGlyHisGlnProA 33
 |||||
 262 AGGGAGGCGCTTGGCCCTCCCTTTGGATGCTGTAATCTACTGTGTCCTA 311
 |||||
 33 lAcGlyAsnArgArgGly...GluAlaArgArgLeuAlaProAsnPheArg 48
 |||||
 312 GAGGAATCTAGTCGGCTTCCGCTTAGCCAGCCCATCTCTCAGTATAGA 361
 |||||
 49 TrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAs 65
 |||||
 362 TGGGATATGATGATCATAGCTTGGAGAACCCACAGCAGCATGAGAGAAGA 411
 |||||
 65 pAspMetGluIlePheMetGluGluMetArgGluIleArgArgGlyLeuA 82
 |||||
 412 GAATATGGAAAGGATGGGAGGGGGTGAGCAGCTGATGGAAAGCTGA 461
 |||||
 82 TgGluLeuGlnLeuArgAnCysLeuArgIleLeuMetGlyGluLeuSer 98
 |||||
 462 GGGAAAGCAGTTGATGATCATAGCTTGGCGGAGTCAGCAGCTGAC...CCC 508
 |||||
 99 AsnHisHisAspHisHisAspGluPheCysLeuMetPro 111
 |||||
 509 CTCACCATGACCATCATATGATGATTTGCTTATGCC 547
 |||||

seq_name: /STD2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810
seq_documentation_block:
 ID AAH75810 standard; cDNA; 767 BP.
 XX AAH75810;
 AC AAH75810;
 DT 17-OCT-2001 (first entry)
 DE Human X chromosome linked gene expression protein 14 coding sequence.
 KW Human; X chromosome linked gene expression protein 14; cancer;

Align seg 1/1 to: AAF59611 from: 1 to: 862

```
3 AsnIleHisGlnGluAsn.....GluGluMetGluGlnProMetG1 16
|||||.....|
271 AATGTCAACAGGAATAATGATGAAAGAGCAAGCAAGTGTCTAA 320
16 nAnGlyGluGlu...AspArgProLeuGlyGlyGlyGlyHisGlnP 32
|||||.....|
321 TAAAGGGAGGAGCCCTGGCCCTACTTTGAATGATGATGCTGCTGC 370
32 foAlaGlyAspArgGly...GlnAlaArgLeuAlaProAsnPh 47
|||||.....|
371 CTAGAGGAAACCTAGGCGGCTTGGAGGAGCCAGCAGCTGCGAGTAT 420
48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyAspG1 64
|||||.....|
421 AGATGGACATAATGATAGCTTGGAGAGCCAGCAGGAGGATGAGAGA 470
64 YAspAspMetGluIlePheMetGluGluMetArgGluIleArgGlyL 81
|||||.....|
471 GAGATATGGAAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 520
81 euArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGlu 97
|||||.....|
521 TGAGGGAAGAGCTTGGATGATGATGATGATGATGATGATGATGAT 567
98 SerAanHisHisAspHisHisAspGluPheCysLeuMetPro 111
|||||.....|
568 CCCCTCCACATGACCATGACGATGATGATGATGATGATGATGATG 609
seq_name: /SIDS2/gcgdata/geneseq/NA2001.DAT:AAF59611
seq_documentation_block:
XX AAF59611 standard: cDNA; 898 BP.
AC AAF59611:
XX
XX
XX 24-APR-2001 (first entry)
XX
XX Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO.76.
XX
XX Cell cycle and proliferation protein; CCYPR; human; agonist;
KW antagonist; gene therapy; detection; gene therapy;
KW transgenic animal disease model; immune disorder;
KW developmental disorder; cell signaling disorder;
KW arteriosclerosis; asthma; allergy; diabetes mellitus;
KW menstrual cycle disorder; bacterial infection; ss.
XX
XX Homo sapiens.
XX
XX W020010741-A2.
XX
XX 01-FEB-2001.
XX
XX 21-JUL-2000; 2000WO-US19948.
XX
XX 21-JUL-1999; 99US-0145075.
XX
XX 08-SEP-1999; 99US-0153129.
XX
XX 10-NOV-1999; 99US-0164647.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
XX
XX MPI: 2001-112727/12.
XX
XX P-FSD8; AAB6047.
XX
XX Human cell cycle and proliferation proteins and polynucleotides are
PT used to treat, diagnose and prevent immune, developmental and cell
PT signalling disorders and cell proliferative disorders including cancer -
XX
```

```
PS Claim 5; Page 181-182; 205pp; English.
XX
XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
CC cell cycle and proliferation proteins (CCYPR), while CCYPR
CC antagonists are used to treat diseases or conditions associated with
CC overexpression of functional CCYPR. Antisense oligonucleotide methods
CC to downregulate expression of functional CCYPR. Antisense oligonucleotide
CC to downregulate expression of functional CCYPR. Antisense oligonucleotide
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using
CC proteins, nucleic acids, agonists or antagonists of human
CC development including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.
XX
XX Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;
SQ
alignment_scores:
Quality: 175.00 Length: 114
Ref: 2.458 Gaps: 4
Percent Similarity: 62.281 Percent identity: 41.228
alignment_block:
US-09-327-750D-13 x AAF59611
Align seg 1/1 to: AAF59611 from: 1 to: 898
3 AsnIleHisGlnGluAsn.....GluGluMetGluGlnProMetG1 16
|||||.....|
299 AATGTCAACAGGAATAATGATGAAAGAGCAAGCAAGTGTCTAA 348
16 nAnGlyGluGlu...AspArgProLeuGlyGlyGlyGlyHisGlnP 32
|||||.....|
349 TAAAGGGAGGAGCCCTGGCCCTACTTTGAATGATGATGCTGCTGC 398
32 foAlaGlyAsnArgArgGly...GlnAlaArgArgLeuAlaProAsnPh 47
|||||.....|
399 CTAGAGGAAACCTAGGCGGCTTGGAGGAGCCAGCAGCTGCGAGTAT 448
48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyAspG1 64
|||||.....|
449 AGATGGACATAATGATAGCTTGGAGAGCCAGCAGGAGGATGAGAGA 498
64 YAspAspMetGluIlePheMetGluGluMetArgGluIleArgGlyL 81
|||||.....|
499 GAGATATGGAAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 548
81 euArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGlu 97
|||||.....|
549 TGAGGGAAGAGCTTGGATGATGATGATGATGATGATGATGATGAT 595
98 SerAanHisHisAspHisHisAspGluPheCysLeuMetPro 111
|||||.....|
596 CCCCTCCACATGACCATGACGATGATGATGATGATGATGATGATG 637
seq_name: /SIDS2/gcgdata/geneseq/NA2000.DAT:AA03880
seq_documentation_block:
ID AA03880 standard: cDNA; 662 BP.
XX
XX AA03880:
XX
```

```
US-09-327-750d-13 x AAC10889
Align seg 1/1 to: AAC10889 from: 1 to: 692
3 AsnIleHisGlnGluAsnGlnGluMetGlnProMetGlnAsnGlyG1 19
||||:||||:||||:||||:||||:||||:||||:||||:||||:
282 AATGTCACACCGAGAAATGATGAAAGAGCAAGTGGCTAA 331
19 uGluAspArgProLeuGly.....GlyGlyGluGlyHisGlnP 32
||||:||||:||||:||||:||||:||||:||||:||||:||||:
332 TAAAGGAGCCCTTGCCCTACCTTGTGATGTTGTGATATCTGTGC 381
32 ToAlaGlyAsnArgArgGly..GlnAlaArgGluAlaProAsnPhe 47
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
382 CTAGAGGAATCTAGGCGGTCCGCTTAGGACGCCATCCTGCAGTAT 431
48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspG1 64
||||| :||| :||| :||| :||| :||| :||| :||| :|||
432 AGATGGGATATGATGATAGCTTGGAGAACACAGGAGGATGAGAGA 481
64 YAspAspMetGluIlePheMetGluMetGluMetArgGluIleArgLysL 81
||||:||||:||||:||||:||||:||||:||||:||||:||||:
482 AGAGATATGGAAAGGATGGGGAGGAGGTGGAGACCTGATGGAAAGC 531
81 euArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeu 97
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
532 TGAGGGAAGAGCAGTTGATCATAGTCTCGCGGCGAGTCAGCATGAC... 578
98 SerAsnHisHisAspHisAspGluPheCysLeuMetPro 111
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
579 CCCCCCTCACCATGACCATCATGATGAGTTTTCNNWATGCCCC 620
```

/csgn2.6/ptodata/2/lna/5b_comb.seq:US-08-487-595-8 -	76.50	120.21	47.95	5073
/csgn2.6/ptodata/2/lna/6a_comb.seq:US-09-024-0208-8	76.50	117.25	70.14	4362
/csgn2.6/ptodata/2/lna/6a_comb.seq:US-09-356-952-12 -	76.50	98.66	76.102	6826
/csgn2.6/ptodata/2/lna/5a_comb.seq:US-08-415-751-47 -	76.00	134.65	7.53	1081
/csgn2.6/ptodata/2/lna/6a_comb.seq:US-09-025-681-42	76.00	132.90	9.42	1294

seq_name: /cgn2_6/ptodata/2/1na/pctus_comb.seq:pct-US92-06840-1

```
seq_documentation_block:
? Sequence 1. Application PC/TUS9206840
? GENERAL INFORMATION:
? APPLICANT : Shell, Inc.
? INVENTOR : Stokely, Edward
? APPLICANT : Shell, Thomas
? TITLE OF INVENTION: Y11 TRANSCRIPTION FACTOR AND METHODS OF
? IDENTIFICATION OF TRANSCRIPTION FACTORS
? TITLE OF INVENTION: ISOLATING SAME
? ADDRESS OF THE APPLICANT: SHELL OIL COMPANY
? CORRESPONDENCE ADDRESS: 10
```

```
seq_documentation_block:
: Sequence 1. Application PC/TUS9206840
```

alignment_scores:		
Quality:	84.50	Length: 86
Ratio:	1.798	Gaps: 3
Percent Similarity:	54.651	Percent Identity: 29.070

```
528 CGGCGACCGCGCTG.....GAACATCGACCGG 553
53 laile.....ProAsnArg...GlnMetAsnAspGlyLeu...GlyGly 75
554 CAGTCGACCGCGCTGCTGCTACCTCGACACAGCGCTCGAAGCGGA 603
76 AspGlyAspAspMetGlnMetPheMetGlnGluMetArg..... 88
604 GATCAACACAGATCAAGAAGTCGGGGGGCGCCCAACGCGCGCTCGGA 653
89 .....GluileArgArgLysLeu 95
654 TCGAGCGCACCGCGCGGAGTTTCAACAGCTGATCTCGCGCGGACGG 703
95 rgGluLeuGlnLeuArgAsnCysLeu...Argile..... 105
704 ATCTACTCCACACGCGAGCGTCTGCTCGGCTTCAACGTCGCGAGCG 753
105 ..... 105
754 CAGCACTACTACTCTCTGACCGCGCGCTGACGCGCGCGGACCA 803
106 LeuMetGlyGlnLeuSerAsnHisHisasp...HisHisasp 118
804 CCTGGTGGCGCACTCTGCGCGCGCGCACCGCTGCTCGCGACGAC 846
seq_name: /cgn2_6/prodata/2/lna/6A_COMB.seq:us-09-171-969-1
seq_documentation block:
; Sequence 1, Application US/09171969
; Patent No. 6284533
; GENERAL INFORMATION:
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: PLASMID-BASED VACCINE FOR TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 75 State Street, Suite 2300
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1807
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; FILING DATE: 01 May 1997 (01.05.97)
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/640,713
; FILING DATE: 01 May 1996 (01.05.96)
; PROSECUTION DATA:
; APPLICATION NUMBER: 08/802,967
; FILING DATE: 21 February 1997 (21.02.97)
; ATTORNEY/AGENT INFORMATION:
; NAME: Leon R. Yankwich
; REGISTRATION NUMBER: 30,237
; REFERENCE/DOCKET NUMBER: TCS 414.1 PCT (05872)
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL:
; ANTI-SENSE:
; FEATURE: Structural coding sequence for
; FEATURE: mature rabbit CMTF
```

```
NAME/KEY:
LOCATION:
PUBLICATION INFORMATION:
AUTHORS: Nagashima, Mariko, et al.
TITLE: Cloning and mRNA tissue distribution of rabbit cholesteryl ester tr
JOURNAL: J. Lipid Res.
VOLUME: 29
ISSUES: 1643 - 1649
PAGES: 1988
DATE: 1988
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1488
US-09-171-969-1
alignment_scores:
Quality: 81.00 Length: 101
Ratio: 1.373 Gaps: 3
Percent Similarity: 58.416 Percent Identity: 28.713
alignment_block:
US-09-327-750d-12 x US-09-171-969-1
Align seg 1/1 to: US-09-171-969-1 from: 1 to: 1488
5 HisGlnGluAsnGluGluMetGluGlnProLeuGlnAsnGlyGluGluAs 21
30 CATCGTGTGTCGTCATCAACCAAGCCGCCCTCTT.....GGTGTGAACC 73
21 PArgProValGlyGlyGlyGluGlyHisGlnProAlaGlyAsnAsnAsnA 38
74 AAGACGCGCCAAAGGTGTGTCACAGCGCTTCCACGCCCGCGCTATCCG 123
38 snAsnAsnHisAsnHisAsnHisAsnHisAsnHisAsnHisAsnHisA 54
124 GACGTCACGCGCGGAGCGGTGTGCTCTCGCGCGGGTCAAGTACGG 173
55 ArgLeuAlaProAsnPheArgTTPAlaIleProAsnArgGlnMetAsnA 71
174 CTTGCACACCTTCAGATCAGCCACCTGTC...CATGCCAGCAGCCAGG 220
71 PGLyLeuGlyGlyAspGlyAspMetGluMetPheMetGluGluMeta 88
221 TGGAGCTGTGGAGCCAGACCATCGACGTCGCCATCCAGACAGTGTC 270
88 rgGluLeuArgArgGlyLeuArgGlnLeuArgAsnGlyLeuArg 104
271 GTGGTCTTCAAGGGGAC...CCTGAAGTACAGCTACACAGTCTCTGGG 317
105 Ile 106
318 GTT 320
seq_name: /cgn2_6/prodata/2/lna/6A_COMB.seq:us-09-036-987A-1
seq_documentation block:
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Saltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Casford, Kathryn P.
; APPLICANT: Maddux, Kristinmarthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Astron Sciences LLC
; STREET: 9330 24th Avenue Road
; CITY: Indianapolis
```

```

95   rglUleuGlnLeu.....ArgAsnCysLeu.....Arg 104
280623 AGCAGCGCAGCGTTCGGCCATCTGTCGGTGGCGGACGACGACCGC 280672

105 IleLeuMetGlyGluLeuSerAsnHis 116
...:|||||:
280673 TTGCACACAGGGGAGTTCATGCGATCATCCGCCAGGATCAC 280714

seq_name : cgn2_6/pdata2/2/na/6B_COMB.seq:US-09-103-840A-1
seq_documentation_block:
  Sequence 1 Identification US/09103840A
  Patent No. 6294328
  GENERAL INFORMATION:
  APPLICANT: FLEISCHMAN, Robert D.
  APPLICANT: WHITE, Owen R.
  APPLICANT: FRASER, Claire M.
  APPLICANT: VENTER, John C.
  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIA
  FILE REFERENCE NO.: TUBERCULOSIS
  CURRENT APPLICATION NUMBER: US/09/103_840A
  CURRENT FILING DATE: 1998-06-24
  NUMBER OF SEQ ID NOS: 2
  SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 1
  LENGTH: 4411529
  TYPE: DNA
  ORGANISM: Mycobacterium tuberculosis
  OTHER INFORMATION: H3/KV
US-09-103-840A-1

alignment_scores:
  Quality: 80.00      Length: 164
  Ratio: 1.013       Gaps: 8
  Percent Similarity: 48.171  Percent Identity: 23.171

alignment_block:
US-09-327-750D-1 x US-09-103-840A-1 ..
Align seg 1/1 to: US-09-103-840A-1 from: 1 to: 4411529

5 HisGlnGuaGlnGluMetGluGlnProLeuGlnAsn...GlyGluG 20
280110 CATCCGCGCAGGAGAACTTGCAGCAAAATCGTCTAAATGTCGGCGTGA 280159
20 uhsApGProValGlyGlyGlyGlyGly.....HisGlnPro 33
280160 CGATGACGCTGTGGCGGATCGAGATGCGCGTGTCTGCACCCG 280209
33 IaGlyArgAsnAsnAsnAsnAsnHis.....AsnHisAsnHisAsn 46
280210 CCGCTGACGCGCGGTGCTGACGCGGCGCTGCAGATCGCGCGGACGATAC 280259
47 HIsHisAArgGlyGlnAlaArg..... 54
280260 GATGACAGGTGCTCTCCGCCAGCTCGACACGCTGCGTGTGATCTCGT 280309
55 .....ATGLeuAlaProAsnPha 61
280310 TACCGCGCATAGACGCCACCGATTGGCGGCGCTGCTCTCGGTGACG 280359
61 TgTrpAlaIleProAsnArgGlnMetAsnArgGlyLeuGlyGly..... 75
280360 GTGGCCGCCGCCACCCCGGGGATCAGCGAGATGCTCTGCAGCGTCTCCGA 280409
76 .....AspGlyAs 78
280410 GCTATACAGCAACGCTGCTGGAACAGCGCTGCGGAGCCGCTCGGCGA 280459
78 pAspMetGluMetPheMetGluGluMetArgGluLeuIleArgGlyLeu 95
|||||:

```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 663..3164
; us-08-396-001-3

alignment_scores:
  Quality: 79.00      Length: 92
  Ratio: 1.362       Gaps: 5
  Percent Similarity: 63.043  Percent Identity: 26.087

alignment_block:
  US-09-327-750D-12 x US-08-396-001-3 ..
  Align seg 1/1 to: US-08-396-001-3 from: 1 to: 3455

17 ASnglyGluGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAl 33
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2451 AACACGAAATATAGAACCC.....CATACAAAAA 2482

33 eGlyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArgA 50
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2483 TAGTCATATATCATATCATATCATATCATATCATATCATATCATAT 2532

50 rGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIlePro... 65
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2533 ATACATATATATCAATATATATATATATATATATATATATATATAT 2582

66 .....AsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspAs 79
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2583 AATGCTTACCATAGAGAGTACAGCTCTGTACCAATAT..... 2624

79 pMetGluMetPheMetGluGlu..MetArgGluIleArgArgGlyLeuArg 95
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2625 .....TTCTCAACCAATATG.....CACAGATCAGA 2652

96 GluLeuGlnLeuArgAsnCysLeu 103
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2653 AAATTCACCTCTCCGCAACAAATTA 2676

seq_name: /cgn2_6/ptodata/2/lna/68_COMB.seq:us-09-323-433A-3

seq_documentation_block:
; Sequence 3, Application US/09323433A
; Patent No. 6218512
; GENERAL INFORMATION:
; APPLICANT: Hoffmann, Stephen L.
; APPLICANT: Hoffmann, Richard C.
; APPLICANT: Sedegah, Martha
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
; TITLE OF INVENTION: AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
; TITLE OF INVENTION: DELIVERING POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Naval Medical Res. & Dev. Cnd.
; STREET: Bldg. 1, T-12 8901 Wisconsin Ave.
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/155.888
; APPLICATION NUMBER: US/08/155.888
; PCT NUMBER: 0050.1491-003
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, A. David
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: N.C. 75,851
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 295-6759
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 663..3164
; us-08-396-001-3

alignment_scores:
  Quality: 79.00      Length: 92
  Ratio: 1.362       Gaps: 5
  Percent Similarity: 63.043  Percent Identity: 26.087

alignment_block:
  US-09-327-750D-12 x US-09-323-433A-3 ..
  Align seg 1/1 to: US-09-323-433A-3 from: 1 to: 3455

17 ASnglyGluGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAl 33
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2451 AACACGAAATATAGAACCC.....CATACAAAAA 2482

33 eGlyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArgA 50
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2483 TAGTCATATATCATATCATATCATATCATATCATATCATATCATAT 2532

50 rGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIlePro... 65
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2533 ATACATATATATCAATATATATATATATATATATATATATATATAT 2582

66 .....AsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspAs 79
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2583 AATGCTTACCATAGAGAGTACAGCTCTGTACCAATAT..... 2624

79 pMetGluMetPheMetGluGlu..MetArgGluIleArgArgGlyLeuArg 95
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2625 .....TTCTCAACCAATATG.....CACAGATCAGA 2652

96 GluLeuGlnLeuArgAsnCysLeu 103
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2653 AAATTCACCTCTCCGCAACAAATTA 2676

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-155-888-1

seq_documentation_block:
; Sequence 1, Application US/08155888
; Patent No. 6068243
; GENERAL INFORMATION:
; APPLICANT: Hoffmann, Stephen L.
; APPLICANT: Hoffmann, Richard C.
; APPLICANT: Sedegah, Martha
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
; TITLE OF INVENTION: AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
; TITLE OF INVENTION: DELIVERING POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Naval Medical Res. & Dev. Cnd.
; STREET: Bldg. 1, T-12 8901 Wisconsin Ave.
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/155.888
; APPLICATION NUMBER: US/08/155.888
; PCT NUMBER: 0050.1491-003
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, A. David
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: N.C. 75,851
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 295-6759
```


553 TATATGCGCTTCAAACTCCTGAACCTGACCAACCATATGAC..... 594

118 pCupPheCysLeuMetPro 124

|||||.....|

595 ...TTTGGCTCATACCT 609

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAA87147

seq_documentation_block:

ID AAA87147 standard; DNA; 187 BP.

XX AAA87147;

XX 08-JAN-2001 (first entry)

XX Rat hepatocyte carcinogenesis biomarker nucleic acid SEQ ID NO.71.

XX Rat; phenobarbital; carcinogenesis marker; carcinogenesis; detection;

XX Identification; carcinogenic; probe; primer; ds.

XX Rattus norvegicus.

XX WO200044902-A2.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US00503.

XX 29-JAN-1999; 99US-0118078.

XX (SEAR) SEARLE & CO G D.

XX Bunch RT, Curtis SW, Rodi CP, Morris DL;

XX WPI: 2000-505977/45.

XX New nucleic acid encoding a carcinogenic biomarker, induced by

XX phenobarbital treatment of rat hepatocytes, useful for identifying

XX carcinogenic compounds.

XX Claim 1; Page 73; 240pp: English.

XX AAA87080 to AAA87656 represent nucleic acid sequences (NI) encoding a

XX carcinogenesis biomarkers. The carcinogenesis biomarkers are induced by

XX phenobarbital treatment of rat hepatocytes. The nucleic acid sequences

XX can be used to identify carcinogenic compounds. The nucleic acid molecules

XX can be used to derive probes and/or primers for detecting or inducing

XX carcinogenesis, respectively.

XX Sequence 187 BP; 39 A; 48 C; 40 G; 60 T; 0 other;

alignment_scores:

Quality: 193.00 Length: 35

Ratio: 5.514

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750d-12 x AAA87147

Align seg 1/1 to: AAA87147 from: 1 to: 187

90 lLeAArgLysLeuArgLysLeuArgLysLeuArgLysLeuArgLysLeu 106

2 ATCCGGAAGACCTTAGGAGCTGCGCTGAGGAGTGTGCGGTATCT 51

106 uMetGlyGluLeuSerAsnHisHisAspHisHisAspGluPheCysLeu 123

52 TATGGGGAGCTCTCTAATCACCAGCACCATCAGATGAATTCGCTTA 101

123 ePro 124

|||||

102 TGCCCT 106

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:

ID AAH45143 standard; cDNA; 792 BP.

XX AAH45143;

XX 07-SEP-2001 (first entry)

XX Human brain expressed X-linked protein, hHex, coding sequence.

XX Human; brain expressed X-linked protein; cytostatic; auditory; nontropic;

XX hHex; dysmryoplasia; hereditary disease; cancer; tumour; deafness;

XX X-chromosome-binding mental retardation; lissencephalous disease; ss.

XX Homo sapiens.

XX WO200140286-A1.

XX 07-JUN-2001.

XX 27-NOV-2000; 2000WO-CN00502.

XX 30-NOV-1999; 99CN-0124179.

XX (BIOR-) BIORAD GENE DEV LTD SHANGHAI.

XX Mao Y, Xie Y;

XX WPI: 2001-397944/42.

XX P-PSDB; AA89224.

XX Isolated human brain-expressed X-linked polypeptide used to diagnose

XX and treat of dysmryoplasia, hereditary diseases, cancer, tumor,

XX deafness and X-chromosome-binding mental retardation.

XX Claim 5; Page 22; 30pp: Chinese.

XX The present sequence is the coding sequence for a human brain-expressed

XX X-linked protein (hHex). hHex and its coding sequence are useful in the

XX diagnosis and treatment of dysmryoplasia, hereditary diseases, cancer,

XX tumors, deafness, X-chromosome-binding mental retardation and

XX lissencephalous disease. hHex is suitable for identifying

XX hHex coding sequence can be used as primers or probes, or in producing

XX gene chips or microarrays.

XX Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

alignment_scores:

Quality: 184.00 Length: 131

Ratio: 2.452

Percent Similarity: 57.252 Percent Identity: 36.641

alignment_block:

US-09-327-750d-12 x AAH45143

Align seg 1/1 to: AAH45143 from: 1 to: 792

1 MetaLAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17

212 ATGGAATGTCACCAACCAAGAAATGAGAAAGGAGCANGTCTCTAATAA 261

17 RGLYGLUGLU...AspArgProValGlyGlyGlyGlyHisGlnProA 33

262 AGCGAGCCCTTGGCCCTCTTGGATGCTGTGAATCTGTGTGCTTA 311

33 lacGlyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisHisArg 49

312 GAGGAAT.....CGTAGG 325

AAFS9611	standard; CDNA; 898 BP.
AAFS9611;	
24-APR-2001	(first entry)
Human	cell cycle and proliferation protein CYPB-22 cDNA, SEQ ID NO:76.
Cell cycle and proliferation protein; CYPB; human; agonist;	
antagonist; general enzyme model; immunologic disorder;	
developmental disorder; cell signalling disorder;	
cell proliferative disorder; cancer; tumour; anaemia;	
arteriosclerosis; asthma; allergy; diabetes mellitus;	
menstrual cycle disorder; bacterial infection; ss.	
Homo sapiens.	
WO200107471-A2.	
01-FEB-2001.	
21-JUL-2000; 2000WO-US19948.	
21-JUL-1999; 99US-0145075.	
08-SEP-1999; 99US-0153129.	
10-NOV-1999; 99US-0164647.	
(INCY-) INCYTE GENOMICS INC.	
Hillman JL, Lal P, Tang YT, Yue H, Au-Yang J, Bandman O;	
Almazal Y, Yang J, Lu DM, Baugin MK, Patterson C, Shah P;	
WPI: 2001-112727/12.	
P-PSDB: AAB50474.	

```

XX      Sequence 998 BP; 250 A; 186 C; 251 G; 211 T; 0 other;
SQ
alignment_scores:
      Quality: 171.50      Length: 132
      Ratio: 2.287      Gaps: 6
      Percent Similarity: 56.818      Percent Identity: 36.364

```

CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.

XX Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 other;
SQ

alignment_scored:
Quality: 162.50 Length: 126
Ratio: 2.355 Gaps: 4
Percent Similarity: 54.762 Percent Identity: 35.714

alignment_block:

US-09-327-750D-12 x AAC03880 ..

Align seg 1/1 to: AAC03880 from: 1 to: 662

```
3 AsnValHisGlnGluAsn.....GluGluMetGluGlnProLeuGln 16
|||||.....
252 ATGTCACCCAGGAGAAATGATGAAAGAGATGAAAGGAGCAGTGGCTAA 301
16 nAspGlyGluGlu...AspArgProValGlyGlyGlyGlyHisGlnP 32
|||||.....
302 TAAAGGGAGCCCTTGGCCCTACCTTTGATGVTGTGTAATACTGTGTC 351
32 roAlaGlyAsnAsnAsnAsnAsnHisAsnHisAsnHisHis 48
|| |||||
352 CTAGAGGAAT.....CGT 365
49 ArgArgGlyGlnAlaArgLeuAlaProAsnPheArgProAlaIlePr 65
|||||.....
366 AGCGGTCCGGCTTAGCGCCATCCTCGAGTATGATGGATATGAT 415
65 oAsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluM 82
|||||
416 GCATAGCTTGGAGACCAACGAGGAGGATGANAAGAGAGATATGGA 465
82 etPheMetGluGluMetArgLileArgArgLysLeuArgGluLeuGln 98
||| |||||..... |||
466 GGATTGGGAGAGGTGAGACGCTGATGTAAGAGCTGAGGAGAAAGCAG 515
99 LeuArgPheCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHis 115
|||||.....
516 TTGACTCATAGTCTGGCGGAGTGCAGCACTGAC...CCCCCTCACCATGA 562
115 pHisHisAspGluPheCysLeuMetPro 124
|||||.....
563 CCATCATGATGAGTTTGCNNWATGCC 590
```

[illegible]

```

ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE     1 (bases 1 to 793)
AUTHORS      Streusberg, A.
TITLE        Direct Subcloning of the P53 Gene
JOURNAL      J Biol Chem 275:11588-11592 (2000)
NBI-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: csaeb@remail.nih.gov
Tissue Procurement: TCC
Library Accession: 100000000
DNA Library Arrayed by: Rubin Laboratory
Conc Library: 100000000
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Gairaud, Rene Guin,
Leticia Hsiao, Martin Kitzler, Kelli Mathewson, Candice McLeavy, Steven
New, Pavan Pandoh, Anna-Liisa Prabhu, Lorraine Saesdi, Jacqueline
Schlein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Marzia van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium URL at: http://image.llnl.gov
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7657043.
Location/Qualifiers
1..793
/organism="Homo sapiens"
/db_xref="LocusID:27018"
/db_xref="taxon:9606:3357965"
/tissue="MGC"
/tissue_type="Eye,retinoblastoma"
/clone.lib="NR_MGC_16"
/lab.host="DH10B-R"
/note="Vector: pOTB7"
196..531
/codon.start=1
/product="p53MR-associated cell death executor: ovarian
granulosa cell tumor suppressor protein (p53MD)"
/db_xref="db_xref=AA003190.1"
/db_xref="GI:13112031"
/translation="MANIRHSENEEOPMONEEDRPLAGGEGPAGNRRGQARRA
NFWPAINPQINOGMGDDGDMFEEMRIRKLRLQLRNLRLMGLNELSHNSH
HDFECLMP"

BASE COUNT    225 a 170 c 205 g 193 t
alignment_block: 225 a 170 c 205 g 193 t
US-09-327-7500-12 x BC003190 ..

alignment_scores:
Quality: 576.50 Length: 124
Ratio: 5.194 Gaps: 1
Percent Similarity: 89.516 Percent Identity: 84.677

Align seg 1/1 to: BC003190 from: 1 to: 793
1 MetAaAnValHisGInGuaSngInGluMetGluGInProLeuGInAs 17
196 ATGCGCAATATTCACGAGAACGAAAGAGATGAGCGCTATGCGAA 245
17 nclGluGlnuAspGProValGluGlyGluGluGluGlnProLeuG 34
246 TGGAGTGAATATTCATCTCTTGGAGGAGAGTGAAGGCCACGAGCTGCA 295

```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..891
/genre="NAD2"
312..647
/feature="NAD2"
/feature="Involved in the common neurotrophin receptor
p75NTR-mediated signal transduction"
/codon_start=1
/product="p75NTR-associated cell death executor"
/protein_id="AAF75129.1"
/db_xref="GI:8452894"
/translation="MANHSENEEPMQNGEEDPLGGEGHQPAGNRGQARRLA
PNFRAIPNROINDMGDDDDMEIFEMREIRKLRELQNLCLIMGELSNNHD
RHDFCLMP"
BASE COUNT 251 a 182 c 224 g 234 t
ORIGIN

alignment_scores:
  Quality: 576.50      Length: 124
  Ratio: 5.194        Gaps: 1
Percent Similarity: 89.516 Percent Identity: 84.677

alignment_block:
US-09-327-750D-12 x AF187064 ..
Align seg 1/1 to: AF187064 from: 1 to: 891
1 MetAlaAnValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAs 17
|||||
312 ATGGCAATATTACCGAAGAACGACAGACATCGACACCTATCGCAGA 361
17 nGlyGluAsnArgProValGlyGlyGlyGlyHisGlnProAlaG 34
|||||
362 TGGAGGAGGACCGCCCTTTGGAGGAGGTGAAGGCCACCACTGCAG 411
34 IyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArg 50
|||||
412 GAAT.....CCAGCG 422
51 GlyGlnAlaArgLeuAlaProAsnPheArgThrPheIleProAsnAr 67
|||||
423 GGACAGGCTCGCGACTGCCCTATTTTCATGGGCCATACCCATAG 472
67 gGlnMetAsnAspGlyGlyGlyGlyAspGlyAspMetGluMetPheM 84
|||||
473 CGATGATCAATGATGGATGGGTGGAGATGGATGATGATGAAATATCA 522
84 eGluGluMetArgIleuIleArgArgLysLeuArgGluLeuGlnLeuArg 100
|||||
523 TGGAGGATGATGAGAAATCAAGAAACTTAGGAGCTGAGTTGAGG 572
101 AsnCysLeuArgIleLeuMetGlyGlyLeuSerAsnHisAspHisH 117
|||||
573 AATTGTCTGCTATCTTATGGGGAGCTCTCTATACCATGACCATCA 622
117 sAspGluPheCysLeuMetPro 124
|||||
623 TGATGAATTTTCCTTATGCTT 644
```

seq_name: gb_sta.G72708

```
seq_documentation_block: 431 bp DNA STS 08-AUG-2001
LOCUS 431 bp STS
DEFINITION 431 bp STS - porcine spleen Sus scrofa STS
ACCESSION G72708.1 GI:15146738
VERSION G72708.1
KEYWORDS STS.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
AUTHORS
  Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
  1 (bases 1 to 421)
  Fahrenkrug S.C., Freking B.A., Rohrer G.A., Smith T.P.L. and
  Keele J.W.
  Single nucleotide polymorphism (SNP) discovery in expressed porcine
  genes
  JOURNAL
  Unpublished (2001)
  COMMENT
  Contact: Freking BA
  USDA, ARS, US Meat Animal Research Center
  PO Box 166, Clay Center, NE 68933-0166, USA
  Tel: 402 762 4278
  Fax: 402 762 4173
  Email: freking@email.marc.usda.gov
  Primer A: GCAAATGGGTGAACCTACT
  Primer B: CCACGACCAATACG
  STS size: 500
  STS Profile:
  ReStart: 95 degrees for 15 minutes
  Denature: 95 degrees for 30 seconds
  Anneal: 56 degrees
  Extension: 68 degrees for 2 minutes
  Cycles: 32 to 45
  Protocol:
  Template: 50-200 ng genomic DNA
  Primer: each 20 pmol
  dNTPs: each 88 uM
  Taq Polymerase: 0.25 units (Qiagen HotStar)
  Buffer: Commercially supplied Qiagen HotStar buffer
```

The STS is derived from PCR amplicons generated from genomic DNA, sequenced from each end using the amplification primers. The sequence does not necessarily represent the entire amplicon. Sequence derived from PolyPhred was trimmed from each end of each unique contig until five consecutive bases exceeded a quality score threshold of 20, and the next 10 bases averaged a quality score of 20 or greater. Amplicon size was estimated by agarose gel electrophoresis.

```
FEATURES
source
  Location/Qualifiers
    .421bp="Sus scrofa"
    /strain="white composite, duroc, meishan, minzhu,
    fengling, crossbreds"
    /db_xref="taxon:9823"
    /sex="male and female"
    /clone_lib="SCF - porcine spleen"
    /dev_stage="adult"
    /note="Organ: spleen"
    <1..421
BASE COUNT 251 a 182 c 224 g 234 t 2 others
ORIGIN
```

```
alignment_scores
  Quality: 411.00      Length: 77
  Ratio: 5.338        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 96.104
```

alignment_block:

US-09-327-750D-12 x G72708/rev ..

Align seg 1/1 to reverse of: G72708 from: 1 to: 421

```
48 HisArgGlyGlnAlaArgLeuAlaProAsnPheArgThrPheIle 64
|||||
421 AATAGACGGGACAAAGCTCGCGACTTGCCTTATTTCCGATGGCCAT 372
64 eProAsnArgGlnMetAsnAspGlyGlyLeuGlyAspGlyAspMetG 81
|||||
371 ACCCAATAGCAGATCAATATGGGATGGGTGGAGATGAGATGATGG 322
```


[illegible]

[illegible]

```

alignment_scores:      Quality: 264.50      Length: 123
                       Ratio: 3.112      Gaps: 4
Percent Similarity: 69.106 Percent Identity: 47.967

alignment_block:
US-09-327-750-12 x HS635G19 ..

Align seg 1/1 to: HS635G19 from: 1 to: 69648

3 AsnValHicGicGluGbaGcGluMetGluGln...ProteuGinAcl 18
6438A ATGCTCCCACTACGATGAAACGAGGAGGAGCCGCCACCCGAGATGA 64387

```

```

1594..16040
/Note="MIR repeat: matches 102..144 of consensus"
repeat_region
16116..16184
/Note="MIR repeat: matches 3..70 of consensus"
repeat_region
16897..17005
/Note="MIR repeat: matches 9..118 of consensus"
repeat_region
17618..17700
/Note="MIR repeat: matches 108..192 of consensus"
repeat_region
17748..17912
/Note="MIR repeat: matches 73..245 of consensus"
repeat_region
18657..18871
/Note="MIR repeat: matches 35..252 of consensus"
repeat_region
18926..19215
/Note="MIR repeat: matches 1..301 of consensus"
repeat_region
19492..19559
/Note="MIR repeat: matches 2..67 of consensus"
repeat_region
19560..19821
/Note="MIR repeat: matches 34..295 of consensus"
repeat_region
19822..19959
/Note="MIR repeat: matches 67..212 of consensus"
repeat_region
19987..20286
/Note="MIR repeat: matches 1..300 of consensus"
repeat_region
20316..20512
/Note="MIR repeat: matches 82..262 of consensus"
repeat_region
20518..20743
/Note="MIR repeat: matches 5970..6224 of consensus"
repeat_region
20769..21074
/Note="MIR repeat: matches 1..303 of consensus"
repeat_region
21085..21207
/Note="MIR repeat: matches 1776..1898 of consensus"
repeat_region
21212..21391
/Note="MIR repeat: matches -6..178 of consensus"
repeat_region
21392..21471
/Note="MIR repeat: matches 3..86 of consensus"
repeat_region
21472..21641
/Note="MIR repeat: matches 1..296 of consensus"
repeat_region
22526..22610
/Note="MIR repeat: matches 267..349 of consensus"
repeat_region
22940..23204
/Note="MIR repeat: matches 378..650 of consensus"
repeat_region
23205..23514
/Note="MIR repeat: matches 1..311 of consensus"
repeat_region
23919..24194
/Note="MIR repeat: matches 36..311 of consensus"
repeat_region
24496..24590
/Note="MIR repeat: matches 253..349 of consensus"
repeat_region
24600..24694
/Note="MIR repeat: matches 102..143 of consensus"
repeat_region
25378..25555
/Note="MIR repeat: matches 2330..2519 of consensus"
repeat_region
25673..25910
/Note="MIR repeat: matches 4..255 of consensus"
repeat_region
26722..26798
/Note="MIR repeat: matches 2572..2649 of consensus"
repeat_region
27040..27107
/Note="MIR repeat: matches 2489..2500 of consensus"
repeat_region
27108..27192
/Note="MIR repeat: matches 41..304 of consensus"
repeat_region
27375..27471
/Note="MIR repeat: matches 2372..2489 of consensus"
repeat_region
27651..27945
/Note="MIR repeat: matches 1..299 of consensus"
repeat_region
28235..28876
/Note="MIR repeat: matches 1113..1817 of consensus"
repeat_region
28975..29097
/Note="MIR repeat: matches 42..173 of consensus"
repeat_region
29173..29196
/Note="MIR repeat: matches 2 mer aa 100 conserved"
misc_feature
29318..30026
/Note="match: GSS: En:AQ075670"
repeat_region
30392..30597
/Note="MIR repeat: matches 35..252 of consensus"
repeat_region
30845..31010

```

```

/Note="MER91B repeat: matches 2..162 of consensus"
repeat_region
31056..31094
/Note="MIR repeat: matches 218..256 of consensus"
repeat_region
31728..31791
/Note="16 copies 4 mer 1979 76 conserved"
repeat_region
31858..31952
/Note="MIR repeat: matches 48..153 of consensus"
misc_feature
complement(33437..33589)
/Note="match: SRS: En:H5731"
repeat_region
33543..33593
/Note="MIR repeat: matches 84..165 of consensus"
repeat_region
33785..33966
/Note="MIR repeat: matches 1..185 of consensus"
alignment_scores
Quality: 240.50 Length: 67
Ratio: 4.076 Gaps: 2
Percent Similarity: 88.060 Percent Identity: 77.612
alignment_block:
US-09-327-750D-12 x HS71487
Align seg 1/1 to: HS71487 from: 1 to: 98274
56 LeuAlaProAspPheArgTIPAlaIleProAsnArgGlnMetAsnAspG1 72
|||||
62887 TTGGCCCTAATTTTCAATGGGCCCATACCCCAATAGCAGGTCATGATGG 628936
72 YLeuGlyGlyAspGlyAspMetGluMetPhenMetGluGluMetArgG 89
|||||
62937 GATGATAGAGTGGAGATGATATGAAATGTCATGGAGGATGAGAG 62986
89 LullehrArg.LysLeuArgGluLeuGlnLeuArgGlnCysLeuArgG1 105
|||||
62987 GATCAGGAGAGAT...AAGGAGTACATGAGGATTCCTGCTAT 63033
105 eLeuMetGlyGluLeuSerAsnHisHisAspHisAspGluPheCys 121
|||||
63034 CCTATGGGAGG.CTGTGTAATCCCATGACCATCATGATGAATTTGC 63081
seq_name: gb_sts.G35294
seq_documentation_block:
LOCUS G35294 STS SHGC-37409, sequence tagged site.
DEFINITION Human, chromosome 10, STS.
VERSION G35294.1 GI:2459462
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 477)
AUTHORS Myers,R.M.
JOURNAL Unpublished (1997)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myerseshgc.stanford.edu
Primer A: ACATCTTTCATGAAGTGTGATG
Primer B: CTTTGGCATCTCTCCAA
STS size: 106
PCR Profile:
Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 15 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds

```

452 CTGAGGGAAGCAGTTGAGTCATAGTCTCGGGGAGTCAGCAGTAC... 499
110 userAsnHisHisAspHisHisAspGluPheCysLeuMetPro 124
|||||
500 .CCCCCTCCACCATGACCATCATGATGAGTTTGCTTTAGGCC 541

LOCUS AF220189 828 bp mRNA PRI 04-MAY-2000
DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA,
complete cds.
ACCESSION AF220189
VERSION AF220189.1 GI:7689028
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. to 828 bp
Xiao, S., H. Gao, G., Ren, S., Chen, Z., and Han, Z.
TITLE A novel gene expressed in human hypothalamus
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 828)
Xiao, S., Song, H., Gao, G., Ren, S., Chen, Z., and Han, Z.
REFERENCE Direct Submission
TITLE Submitted (30-DEC-1999); Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
source 1..828 bp
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
167..544
/codon_start=1
/product="uncharacterized hypothalamus protein HBEX2"
/protein_id="AAF67634.1"
/db_xref="gi:7689028"
/translation="MESKRRVNSLNMENQNEEKQVANGKEPLALPIDAGETC
VPNGRRRVRQPIQTRKRDHRLCEFOARMRENRIGEEVQMLEKLRQLS
HSLAKSIDPHHNDLRFCLP"
BASE COUNT 256 a 170 c 221 g 181 t
ORIGIN

alignment_scores: Quality: 191.00 Length: 131
Ratio: 2.513 Gaps: 5
Percent Similarity: 58.015 Percent Identity: 37.405
alignment_block:
US-09-327-750D-12 x AF220189 ..
Align seg 1/1 to: AF220189 from: 1 to: 828
1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
||| |||||
206 ATGGAAAATGCCAACCAAGAAATGAAGAAAGAGCAAGTTCCTAATAA 255
17 nclygluglu...AspArgProValGlyGlyGlyGlyHisGlnProA 33
:||||| |||||
256 AGGGAGGCCCTTGGCCCTCTTGGATGCTGCTGGAATACATGCTGCTCA 305
||| |||||
33 IagLysAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArg 49
||| |||||
306 GAGGAAT.....CGTAGG 319
50 ArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTIPAlaIle.... 64
||| |||||
320 CGGTTCGCGTTAGCAGCCATCTCTGAGATAGATAGATGATGCA 369
65ProAsnArgGlnMetAsnAspGlyLeuGlyGlyAspG 77
|||||
370 TAGCTTGGAGAACCCAGCAGGATGAGAGAA..... 403
77 IyAspAspMetGluMetPheMetGluMetGluMetArgGlyLys 93
||| |||||
404 ..GAGATATGGAAGGATGGGAGGAGCTGATGATGGAAG 451
94 LeuArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGlu 110
||||| ||||| ||||| |||||


```

VERSION      BE334866.1  GI:9208642
KEYWORDS     EST
SOURCE       house mouse.
ORGANISM     Mus musculus
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        M1-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT      Tumor Gene Index
              Unpublished (1997)
              Email: Roba-strausberg.ph.d.
              This clone is available royalty-free through LLNL; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:1069792
              Seq primer: -40RP from Gibco
              High quality sequence stop: 465.
              Location/Qualifiers
                1..542
                /organism="Mus musculus"
                /db_xref="taxon:10099"
                /clone_lib="Sc23286"
                /sex="female (lactating)"
                /tissue_type="mammary gland"
                /lab_host="DHI08"
                /note="Vector: pT73D-Pac (Pharmacia) with a modified
                polylinker; 1st strand cDNA was prepared from mammary
                gland tissue from a lactating female, and was then primed
                with a Not I - oligo(dT) primer. Double-stranded cDNA was
                ligated to Eco RI adaptors (Pharmacia); digested with Not
                I and cloned into the Not I and Eco RI sites of the
                modified vector. The library is normalized. Library
                was constructed by Benito Soares and M. Fatima Bonaldo."
              BASE COUNT      156 a 141 c 164 g 81 t
              ORIGIN
                alignment_scores:
                  Quality: 688.00      Length: 124
                  Ratio: 5.548          Gaps: 0
                  Percent Similarity: 100.000      Percent Identity: 98.387
                alignment_block:
                  US-09-327-750D-12 x BE334866
                  Align seg 1/1 to: BE334866 from: 1 to: 542
                  1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
                  166 ATGGCCCATGTCCACGAGAAACGAGAGATGGAGCAGCCCTGCAGAA 215
                  17 nGlyGluAsnArgProValGlyGlyGlyGlyHisGlnProAlaG 34
                  216 TGGACAGAGACCCCTGTGGAGAGAGTGGGCGCCACGACCTGCTG 265
                  34 LyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArgArg 50
                  266 CAAACACACACACACACACACACACACACACACACACACACAC 315
                  51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrrpAlaIleProAsnAr 67
                  316 GCCCAGGCTGGCCGACTTCCCTTAACCTCCGATGGGCCATCCCAACAG 365
                  67 GlnMetAsnAspGlyLeuGlyGlyAspGlyAspAspMetGluMetPheM 84
                  366 CGAGATGATGACGGCTGGTGGAGATGGAGATGGATGATGGAAATGTTC 415
                  84 eGluGluMetArgGluLeuArgArgGlyLysLeuArgGluLeuArg 100
                  416 TGGAGGAGATGAGAGATCCCGGAGAAAGCTTAGGAGGTACAGCTGAGA 465

```

```

101 AsnCysLeuArgTleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
117 sAspGluPheCysLeuMetPro 124
516 TGATGAATCTGCCTTATGCCT 537
seq_name: gb_est1:AA215070
seq_documentation_block: 557 bp mRNA EST 03-SEP-1997
LOCUS AA215070.1 Mus musculus embryonic carcinoma (#937317) Mus
DEFINITION musculus cDNA clone IMAGE:551487.5; similar to gb:M38188 OVARIAN
GRANULOSA CELL 13.0 KD PROTEIN HGR74 (HUMAN); mRNA sequence.
ACCESSION AA215070.1 GI:1814831
VERSION AA215070.1
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          (bases 1 to 557)
          Email: Allen.M., Boyles.M., Dietrich.N., Dubuque.T.,
          Geisel.S., Kucaba.T., Lacy.M., Le.M., Martin.J., Morris.M.,
          Schellenberg.K., Steptoe.M., Tan.F., Underwood.K., Moore.B.,
          Theising.B., Wylie.T., Lennon.G., Soares.B., Wilson.R. and
          Waterson.R.
          The WashU-HHMI Mouse EST Project
          Unpublished (1996)
          Contact: Marra M/Mouse EST Project
          WashU-HHMI Mouse EST Project
          Washington University School of MedicineP
          444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: mouseest@wustl.wustl.edu
          This clone is available royalty-free through LLNL; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          MGI:397335
          Seq primer: -28m13 rev1 ET from Amersham
          High quality sequence stop: 447.
          Location/Qualifiers
            1..357
            /organism="Mus musculus"
            /db_xref="taxon:10099"
            /clone_lib="Stratagene mouse embryonic carcinoma (#937317)"
            /tissue_type="carcinoma"
            /dev_stage="embryonic"
            /lab_host="SOLR (kanamycin resistant)"
            /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
            XhoI; Cloned unidirectionally. Primer: Oligo dT. P19 cell
            line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
            adaptor sequence: 5' GATTCGCGACGAG 3' -3' adaptor
            sequence: 5' TCGAGTTTATTTTTTTTTT 3'
          BASE COUNT      159 a 145 c 165 g 88 t
          ORIGIN
            alignment_scores:
              Quality: 688.00      Length: 124
              Ratio: 5.548          Gaps: 0
              Percent Similarity: 100.000      Percent Identity: 98.387
            alignment_block:
              US-09-327-750D-12 x AA215070
              Align seg 1/1 to: AA215070 from: 1 to: 557
              1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17

```



```

polyA_signal      790..795
polyA_site        809
BASE COUNT        208 a 211 c 227 g 163 t
ORIGIN
alignment_scores:
  Quality: 688.00      Length: 124
  Ratio: 5.548         Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 98.387
alignment_block:
  US-09-327-750D-12 x AK004531
  Align seg 1/1 to: AK004531 from: 1 to: 809
1 MetAlaAsnValHisGInGluAsnGluMetGluGInProLeuGlnAs 17
181 ATGGCCAAATGTCACCAAGAAACCAAGAGATGGAGCAGCCCTCGAGAA 230
17 nGlyGluLeuAspArgProValGlyGlyGluGluGlnProAlaG 34
231 TGGCAGAGAGACCGCCCTGTGGAGAGGTGAGGCCACACCTGCTG 280
34 LyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHis 50
281 CAACACACACACACACACACACACACACACACACACACACACAC 330
51 GlyGlnAlaArgLeuMetGluGluLeuSerAsnHisAsnHis 67
331 GCCCAGGCTGCCGCTTATGGGGGAGCTGTACCCACACACACAC 380
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPhe 84
381 GCAGATGATGACGGTGTGGTGGAGATGGAGATGATGGAAATGTTC 430
84 etGluGluMetArgGluLeuArgGlyLysLeuArgGluLeuArg 100
431 TGGAGAGATGAGAGAGATCCGGAAGAAGCTTAGGAGCTACAGCTGA 480
101 AsnCysLeuArgLeuMetGlyGluLeuSerAsnHisAsnHis 117
481 AATTCTCTACGCTTATGGGGGAGCTGTACCCACACACACAC 530
117 saspluPhcCysLeuMetPro 124
531 TGATGAATCTGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 552
seq_name: 9b_est1:AA253897

seq_documentation_block: 468 bp mRNA EST 13-MAR-1997
LOCUS AA253897.1 Barstead mouse pooled organs MRLRB4 Mus musculus cDNA
DEFINITION MRLRB4 IMAGE:592482.5.13.0 KD. PROTEIN HGR74 (HUMAN); mRNA sequence.
ACCESSION AA253897
VERSION AA253897.1 GI:1888253
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1. (bases 1 to 468)
AUTHORS Karim, M., Hilder, T., Allen, M., Bowles, M., Dietrich, M., Dubouque, T.,
Garcia, S., Kasper, T., Lee, M., Le, M., Martin, M., Miller, B.,
Schallenberg, K., Stepien, M., Tan, F., Underwood, K., Wilson, R.,
Thelning, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project

```

WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MDR:431042
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 455.
Location/Qualifiers
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```
101 ascCysLeuArgIleLeuMetGluLeuSerAsnHisAspHis117
|||||
478 AATTGCTACGATCTTATGGGAGCTGCTTAACCAACGATCACCACCA 527
|||||
117 sAspGluPheCysLeuMetPro 124
|||||
528 TGATGAATTCGCTTATGCGCT 549
|||||

seq_name: gb_est1:AI152323
seq_documentation_block: 430 bp mRNA EST
LOCUS AI152323 430 bp mRNA EST 30-SEP-1998
DEFINITION ud79a02.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:1477034 5' similar to gb:h38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN); mRNA sequence.

ACCESSION AI152323
VERSION AI152323.1 GI:3680792
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 430)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tao,F., Underwood,K., Moore,B.,
Theisinger,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterson,R.H.M.I Mouse EST Project
UNPUBLISHED (1996)
JOURNAL
COMMENT Contact: Marra M/Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouse@waterson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 419.
Location/Qualifiers
1. 430
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="1477034"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/lab_host="DH10B"
/notes="Vector: pT773D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT773 vector. Library is normalized. Library
was constructed by Benito Soares and M. Fatima Bonaldo."
BASE COUNT 135 a 111 c 113 g 71 t
ORIGIN
source
1. 430

alignment_scores:
Quality: 681.00 Length: 124
Ratio: 5.537 Gaps: 0
Percent Similarity: 99.194 Percent Identity: 97.581

alignment_block:
US-09-327-750d-12 x AI152323 ..
Align seg 1/1 to: AI152323 from: 1 to: 430
```

```
1 MetAlaAsnValHisGlnGluAsnGluLeuMetGluGlnProLeuGlnAs 17
|||||
54 ATGCCAATGTCCACAGAGAAACGAGAGATGGAGCAGCCCTGCAGAA 103
|||||
17 sGlyGluGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
104 TGGCAGGAGAGACCCCTCTGGATGGAGGTGAGGGCCACAGCCGCTG 153
|||||
34 lYAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisAsnG 50
|||||
154 CAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 203
|||||
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAlaIleProAsnAr 67
|||||
204 GCCCAGGCTCGCGACCTGCCCTTAACCTCCGATGGGCGCATTCGCCA 253
|||||
67 gCtMetAsnAspGlyLeuGlyGlyAspClyAspAspMetGluMetPheM 84
|||||
254 GCAGATGAATGACGGGTGGGTGGAGATGGAGATGATGGAAATGTTCA 303
|||||
84 etGluGluMetArgGluIleArgArgGlyLeuArgGluLeuGlnLeuArg 100
|||||
304 TGGAGGAGATGAGAGATCCCGAGAAAGCTTAGGAGAGCTACAGCTGAG 353
|||||
101 AscCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAspHis117
|||||
354 AATTGCTACGATCTTATGGGAGCTGCTTAACCAACGATCACCACCA 403
|||||
117 sAspGluPheCysLeuMetPro 124
|||||
404 TGATGAATTCGCTTATGCGCT 425
|||||

seq_name: gb_est1:AW908751
seq_documentation_block: 503 bp mRNA EST 25-MAY-2000
LOCUS AW908751 503 bp mRNA EST 25-MAY-2000
DEFINITION IMAGE:1515440 5' similar to gb:h38188 OVARIAN GRANULOSA CELL 13.0
IMAGE:1515440 5' similar to gb:h38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN); mRNA sequence.

ACCESSION AW908751
VERSION AW908751.1 GI:8073984
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 503)
AUTHORS Kaji,C., Kaji,T., Kaji,M., Kaji,N., Kaji,N., Kaji,N., Kaji,N.,
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: uf57a05.x1
Contact: Robert Strausberg, Ph.D.
Email: cgepbs-re@mail.nih.gov
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40RP from Gibco
High quality sequence stop: 468.
Location/Qualifiers
1. 503
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="1515440"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/lab_host="DH10B"
/notes="Vector: pT773D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT773 vector. Library is normalized. Library
was constructed by Benito Soares and M. Fatima Bonaldo."
BASE COUNT 135 a 111 c 113 g 71 t
ORIGIN
source
1. 503
```



```

MEDLINE 20298829
REFERENCE 2 (bases 1 to 891)
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P.,
TITLE Direct Submission, Li,Y., Irie,S. and Sato,T.
JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia
University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA
FEATURES
source 1..891
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1..891
CDS 312
/gene="NADE"
/note="Involved in the common neurotrophin receptor
p75NTR-mediated signal transduction."
/product="p75NTR-associated cell death executor"
/protein_id="AAF75129.1"
/db_xref="GI:8452894"
/translation="MANTHGEEMEPNQGDEPRPGGEGHOPAGNRRQARRLA
PNFRATPNQINDGGDDGDEIFMEIRKRLRELQLNCLRLMGLSLNHH
HDEFLMP"
BASE COUNT 251 a 182 c 224 g 234 t
ORIGIN

alignment_scores:
Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-327-750d-13 x AF187064
Align seg 1/1 to: AF187064 from: 1 to: 891
1 MetAlaAnlleHisGlnGluAsnGluMetGluInProMetGlnAs 17
|||||
312 ATGCCAAATATTCACGAGAAACGAGAGAGATGAGCGCCTATGCAGAA 361
17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
362 TGGAGAGAGACCGCCCTTTGGAGAGAGTGAAGGCCACCGCTGCAG 411
34 IyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTirPaLa 50
|||||
412 GAAATCGACGGGGACAGCGTCGCCGACTTGCCCTTAATTTTCGATGGGCC 461
51 lleProAsnArgGlnIleAsnAspGlyMetGlyIyAspGlyAspAspMe 67
|||||
462 ATACCAATAGCAGATCAATGATGGATGGTGGATGGAGTGGATGATGT 511
67 tGluIlePheMetGluGluMetArgGluIleArgArgIyAspGlyLeuArgGluL 84
|||||
512 GGAATATTATTCAGGAGAGATGAGAGAAATCAGAGAAACTTAGGAGC 561
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
562 TCCAGTTCGAGAAATTCGCTGCTATCTTATGGGAGAGCTCTCTATATC 611
101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
612 CATGACCATCATGATGAATTTTGCCTTATSCCT 644
seq_name: gb_ro:AF187066

seq_documentation_block:
LOCUS AF187066 700 bp mRNA ROD 11-JUN-2000
DEFINITION Mus musculus p75NTR-associated cell death executor (NADE) mRNA,
complete cds.
ACCESSION AF187066
```

```

AF187066.1 GI:8452897
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 700)
AUTHORS Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D.,
Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.
TITLE A p75NTR-associated cell death executor, is involved in
signal transduction mediated by the common neurotrophin receptor
p75NTR.
JOURNAL J Biol Chem. 275 (23), 17566-17570 (2000)
MEDLINE 20298829
REFERENCE 2 (bases 1 to 700)
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P.,
Hanaoka,T., Li,Y., Irie,S. and Sato,T.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia
University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA
FEATURES
source 1..700
/organism="Mus musculus"
/db_xref="taxon:10090"
gene 1..700
CDS 177
/gene="NADE"
/note="Involved in the common neurotrophin receptor
p75NTR-mediated signal transduction. NADE"
/product="p75NTR-associated cell death executor"
/protein_id="AAF75129.1"
/db_xref="GI:8452894"
/translation="MANTHGEEMEPNQGDEPRPGGEGHOPAAANNNNHHNH
HNRHRRQARRLAAPNFRATPNQINDGGDDGDDNDFMEIRKRLRELQLNCL
RLMGLSLNHHHDEFLMP"
BASE COUNT 178 a 187 c 203 g 132 t
ORIGIN

alignment_scores:
Quality: 567.50 Length: 124
Ratio: 5.113 Gaps: 1
Percent Similarity: 89.516 Percent Identity: 83.065

alignment_block:
US-09-327-750d-13 x AF187066
Align seg 1/1 to: AF187066 from: 1 to: 700
1 MetAlaAnlleHisGlnGluAsnGluMetGluInProMetGlnAs 17
|||||
177 ATGGCAATATTCACGAGAAACGAGAGATGAGCGCCTATGCAGAA 226
17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGlnProAlaG 34
|||||
227 TGCACAGAGAGACCGCCCTGTGGGAGAGGTGAGGCCACCGCCTGCTG 276
34 IyAsn.....ArgArg 37
|||||
277 CAACACACACACACACACACACACACACACACACACACACACAGAGA 326
38 GlyGlnAlaArgArgLeuAlaProAsnPheArgTirPaLleIleProAsnAr 54
|||||
327 GGCCAGGCTGCCGACTTGCCCTTAATTTCCGATGGGCGCATTCGCCACAG 376
54 gGlnIleAsnAspGlyMetGlyIyAspGlyAspAspMetGluIlePhe 71
|||||
377 CGAGTGAATGACGGTGTGGAGATGGAGATGATGATGAATGTCA 426
611
71 etGluGluMetArgGluIleArgArgIyAspGlyGluLeuGlnLeuArg 87
```

/translucio="MEOPLOQEDRPVGGEGHOPANNNNHHNNHHNNHHNNH
RGGARLAPNFNFAINPNMNDGLGDDMEMEEMRETRKRLRELQRLNCLRI
MGLSNHHDHDFCLAP"

BASE COUNT 142 a 130 c 136 g 111 t
ORIGIN

alignment_scores:
Quality: 538.50 Length: 125
Gaps: 1
Percent Similarity: 84.800 Percent Identity: 79.200

alignment_block:

US-09-327-750D-13 x AF187065

Align seg 1/1 to: AF187065 from: 1 to: 519

6 GlnGluAsnGluGluMetGluGlnProMetGlnAsnGlyGluLeuSepar 22
|||||
29 CNGSAAACACAGAAATGGACGAGCCCGCGAGATGACACAGAGAGCG 78
|||||
22 gProLeuGlyGlyGlyGluGlyHisGlnProAlaGlyAsn..... 35
|||||
79 CCGCGTGGGAGAGGTGAGGCGCCAGCCGCTGCTGCAACACACACAC 128
36Arg 36
129 ACACCTACACACACACACACACACACACACACACACACATCATCCG 178
37 ArgGlyGlnAlaArgGluAlaProAsnPheArgTpaAlaIleProAs 53
|||||
179 AGAGGACAGSCCGCGCGACTTGCCTTAATTCGATGGCCATTCGCCAA 228
53 nArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMetGluIleP 70
|||||
229 CAGGACGATCATGATGGGTGGGTGGAGATGATGATGATGATGATG 278
70 hMetGluGluMetArgGluIleArgGlyLeuArgGluLeuGluLeu 86
|||||
279 TCATGGAGAGATGAGAGATCGGAGAAAGCTTAGGGAGCTGCAGTTG 328
87 ArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAspH1 103
|||||
329 AGAAATTCCTCGCTATCTTATGGGGGAGCTCTCTAATCACCAGCA 378
103 sHisAspGluPheCysLeuMetPro 111
379 TCACGATGATCTGCCTTAGGCT 403

seq_name: gb_sts:G72708

seq_documentation_block:

LOCUS G72708 421 bp DNA 08-AUG-2001
DEFINITION MARC 4953-4954:991939031:1 SCF - porcine spleen STS
genomic, sequence tagged site.
ACCESSION G72708
VERSION G72708.1 GI:15146738
SOURCE Pigs.

ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE 1 (bases 1 to 421)
AUTHORS Fehrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L. and
Keele,J.W.

TITLE Single nucleotide polymorphism (SNP) discovery in expressed porcine
genes
JOURNAL Unpublished (2001)

COMMENT
Contact: Freking BA
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4278

Fax: 402 762 4173
Email: freking@email.marc.usda.gov
Primer A: GCMAATGGGTGAACCTACT
Primer B: CCACGACCAATAGACG
STS size: 400
PCR Product: 400

Hotstart: 95 degrees for 15 minutes
Denature: 95 degrees for 30 seconds
Anneal: 56 degrees
Extension: 68 degrees for 2 minutes
Cycles: 32 to 45

Protocol:
Template: 50-200 ng genomic DNA
Primer: each 20 pmol
dNTPs: each 80 uM
Req Polymerase: 0.25 units (Qiagen HotStar)

Buffer:

Commercially supplied Qiagen HotStar buffer

The STS is derived from PCR amplicons generated from genomic DNA, sequenced from each end using the amplification primers. The sequence does not necessarily represent the entire amplicon. Each sequence derived from polyphred was trimmed from each end, the unique contig was then constructed. The contig was extended a quality score of 20 or greater. Amplicon size was estimated by agarose gel electrophoresis.

FEATURES

Location/Qualifiers

1..421

/organism="Sus scrofa"
/strain="white composite, duroc, meishan, minzhu,
fengjing, crossbreds
/db_xref="taxon:9823"
/taxonomy="eukaryote"
/dev_stage="adult"
/note="Organ: spleen"

BASE COUNT 130 a 101 c 86 g 102 t 2 others
ORIGIN

alignment_scores:

Quality: 418.00 Length: 77
Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.701

alignment_block:

US-09-327-750D-13 x G72708/rev ..

Align seg 1/1 to reverse of: G72708 from: 1 to: 421

35 AsnArgArgGlyGlnAlaArgGluAlaProAsnPheArgTpaAlaI 51
|||||
421 AATACCGGAGACGCTCCGCGACTTGCCTTAATTCGATGGCCAT 372
51 eProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMetG 68
|||||
371 ACCCATAGGCAGATCAATGATGGATGGGTGGAGATGGAGATATGC 322
68 lullePheMetGluGluMetArgGluIleArgGlyLeuArgGluLeu 84
|||||
321 AATGTTCTATGGAGGATGAGAAATCAGAGAAATTAAGGAGCTG 272
85 GlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 101
|||||
271 CAGTTGAGGAATGTTCTCGCTATCTTATGGGGAGCTCTCTAATCA 222
101 sasPHisHisAspGluPheCysLeuMetPro 111
|||||
221 TGACCATCATGATGAATTTGGCTTATGCTT 191

[illegible]

OM of: US-09-327-750d-13 to: EST: * out_format : pfs

Date: Mar 11, 2002 2:16 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framer_p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09327750/tunat_11032002_101153750/app_query.fasta_1.1472
-DB=EST -PMF=fastp -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPL=0.000 -LOAPEXT=0.500 -CGAPOP=4.500
-GAPEXT=0.100 -GAPOP=10.000 -LOAPEXT=0.500 -CGAPOP=6.000
-GAPEXT=7.000 -GAPOP=10.000 -LOAPEXT=0.500 -DELAP=6.000
-DELAP=7.000 -GAPOP=10.000 -LOAPEXT=0.500 -DELAP=6.000
-ALIGN=5 -CGALIGN=200 -THR_SCORE=100 -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFM=pfs -NORM=ext -HEADSIZE=500
-MILEN=0 -MAXLEN=200000000 -USER=US09327750.ecgn1.1.5654
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPHY -WAIT_THREADS=1

Search information block:

Query: US-09-327-750d-13
Query length: 11
Database: 1151937
Database length: 1077921985
Search time (sec): 4085.940000

score_list:

Sequence	Strd	Orig	25score	5Score	len	Documentation
gb_est2.N34237	+	609	1004.04	3.3e-49	519	1.111237.20-03.ri Soares melanc
gb_est2.BG54378	+	609	1048.04	3.3e-49	535	1.BG54378.1b39e02.y1 NR85 islet
gb_est1.AW512400	+	609	1046.90	3.8e-49	612	1.AW512400.x74b09.x1 NCI_CGAP_Ly
gb_est1.A1193112	+	609	1046.86	3.8e-49	615	1.A1193112.qe69g09.x1 Soares_Fetal
gb_est1.AA576958	+	609	1046.79	3.8e-49	620	1.AA576958.mh2d04.s1 NCI_CGAP_Cc
gb_est1.A1929106	+	609	1046.24	4.1e-49	662	1.A1929106.ue5b10.y1 Schneider_F
gb_est2.BG745809	+	609	1046.08	4.2e-49	674	1.BG745809.6027339171.NIH_MGC_11
gb_est1.AV701506	+	609	1045.53	4.5e-49	720	1.AV701506.AV701506.ADB Homo sap
gb_est2.BF058872	+	600	1031.77	2.6e-48	586	1.BF058872.7A32972.x1 NCI_CGAP_O
gb_est2.BG476585	+	609	1031.77	2.6e-48	586	1.BG476585.7A32972.x1 NCI_CGAP_O
gb_est2.BG476585	+	577	1026.88	4.9e-48	567	1.BG705685.60366891051.NIH_MGC_48
gb_est2.BF896163	+	577	1026.88	4.9e-48	567	1.BG705685.60366891051.NIH_MGC_48
gb_est2.BF896163	+	597	1025.39	6.0e-48	676	1.BE896163.601438756F1.NIH_MGC_72
gb_est2.BG953540	+	596	1025.42	5.9e-48	550	1.BG953540.PMO-HT0911-080201-012
gb_est1.BE252272	+	589	1012.27	7.2e-47	626	1.BE252272.601114115F1.NIH_MGC_16
gb_est1.A1928835	+	585	1005.69	7.5e-47	603	1.A1928835.ue5b10.x1 Schneider_F
gb_est1.AA514324	+	584	1005.92	7.2e-47	479	1.AA514324.df56b04.s1 NCI_CGAP_Cc
gb_est2.BG994124	+	584	1001.81	1.2e-46	779	1.BG746807.60273233F1.NIH_MGC_49
gb_est2.BG746807	+	583	1004.69	8.2e-47	472	1.BG746807.60273233F1.NIH_MGC_49
gb_est2.W23163	+	583	1004.69	8.2e-47	472	1.BG746807.60273233F1.NIH_MGC_49
gb_est2.BF709912	+	583	1003.41	1.0e-46	429	1.BF709912.U13-HP0619-780600-192
gb_est1.BE378166	+	582	1003.41	1.0e-46	429	1.BF709912.U13-HP0619-780600-192
gb_est2.W40428	+	582	1001.99	1.2e-46	507	1.W40428.zb74405.r1 Soares_Fetal
gb_est2.BE935913	+	582	1001.68	1.2e-46	526	1.BE935913.QV2-NM0053-160800-314
gb_est2.BF198483	+	582	1001.67	1.3e-46	527	1.BF198483.248426.MARC.2P1G.Sus
gb_est1.BE255852	+	582	1001.43	1.3e-46	542	1.BE255852.601193356F1.NIH_MGC_7
gb_est1.AW113584	+	582	1001.03	1.4e-46	568	1.AW113584.fcf31901.x1 NCI_CGAP_B
gb_est2.BG773781	+	582	1000.96	1.4e-46	573	1.BG773781.604661455F1.NIH_MGC_21
gb_est1.A1075989	+	582	1000.94	1.4e-46	574	1.A1075989.OV4.60747X1.Scores_Fet
gb_est2.BF715451	+	582	1000.94	1.4e-46	574	1.A1075989.OV4.60747X1.Scores_Fet
gb_est1.AV709328	+	582	1000.57	1.4e-46	600	1.BG715521.602677335F1.NIH_MGC_48
gb_est2.BF709328	+	582	1000.43	1.5e-46	610	1.AV709328.AV709328.ADC Homo sap
gb_est2.BG705415	+	582	1000.36	1.5e-46	615	1.BG705415.602685538F1.NIH_MGC_95
gb_est2.BF750590	+	582	999.81	1.6e-46	656	1.BF750590.bx12a12.y1 Human Iris
gb_est1.AV752634	+	582	999.70	1.6e-46	665	1.AV752634.AV752634.NPD Homo sap
gb_est2.BF183092	+	582	999.65	1.6e-46	669	1.BF183092.601809666F1.NIH_MGC_18
gb_est2.BF032616	+	582	999.54	1.6e-46	678	1.BF032616.601453103F1.NIH_MGC_66
gb_est1.BE774730	+	582	999.49	1.7e-46	682	1.BE774730.601560522F1.NIH_MGC_20
gb_est2.BE898301	+	582	999.44	1.7e-46	686	1.BE898301.601681102F1.NIH_MGC_5
gb_est1.BE729599	+	582	999.44	1.7e-46	686	1.BE729599.601684346F1.NIH_MGC_20

gb_est2.BG925224 + 582.00 999.39 1.7e-46 590 | BG925224.602747641F1.NIH_M

gb_est2.B1111083 + 582.00 999.35 1.7e-46 593 | B1111083.602861185F2.NIH_M

gb_est1.BE385584 + 582.00 999.30 1.7e-46 597 | BE385584.601275870F1.NIH_M

gb_est2.BF970435 + 582.00 999.25 1.7e-46 701 | BF970435.602272771F1.NIH_M

seq_name: gb_est2.N34237

seq_documentation_block:

LOCUS N34237 519 bp mRNA EST 16-JAN-1996
DEFINITION YX79C03.ri Soares melanocyte 2NHRM Homo sapiens cDNA clone
IMAGE:267940 5' similar to gb.N38188 OVARIAN GRANULOSA CELL 13.0 KD
KEYWORDS N34237
ACCESSION N34237.1 GI:1155379
VERSION EST
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillier et al. CO-519. Dubuque, T., Ellington, K., Hawkins, M., Holman,
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, P., Trevaskis, E., Waterston,
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 8820
E-mail: wustl.edu
High quality sequence stops: 449
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: 77
High quality sequence stop: 449.
Location/Qualifiers
1..519
/gene="Homo sapiens"
/gb_xref="taxon:9606"
/clone="IMAGE:267940"
/sex="Male"
/tissue_type="melanocyte"
/lab_host="DH10B (ampicillin resistant)"
/note="Vector: p773D (Pharmacia) with a modified
polylinker. Site 1 Not I. Site 4 Eco RI. Site 15
TGTACCACTGAGTGGAGCGCGCGAGTGTGTGTGT 3'.
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo. RNA from normal foreskin melanocytes
(FS74) was kindly provided by Dr. Anthony P. Albino."

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BASE COUNT 148 a 114 c 153 g 98 t

ORIGIN

alignment_scores:

Quality: 609.00 Length: 111

Ratio: 5.486 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750d-13 x N34237 ..

Align seg 1/1 to: N34237 from: 1 to: 519

DNA Sequencing by: Washington University Genome Sequencing Center
 Cloning through the I.M.G.E. Consortium
 found through the I.M.G.E. Consortium/LLNL at:

Image link: <http://www.ncbi.nlm.nih.gov/ncicgap>

Seq primer: 400P from Gibco

High quality sequence stop: 422.

Location/Qualifiers

1. .612

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:284958"

/clone_lib="MCGAP-Lym12"

Cell="eType="Lymphoma, Follicular mixed small and large

Cell="eType="Lymphoma, Follicular mixed small and large

/lab_host="DH10B"

/note="Organ: lymph node; Vector: pCMV-SPORT6; Site: 1;

SalI; Site: 2: NotI; Cloned unidirectionally. Primer:

Oligo dn. Average insert size 1.25 kb. Life Technologies

catalog #: 11547-015"

BASE COUNT 169 a 145 c 120 g 178 t

ORIGIN

alignment_scores:

Quality: 609.00 Length: 111

Ratio: 5.486 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x AM512400/rev ..

Align seg 1/1 to reverse of: AM512400 from: 1 to: 612

1 MetAlaAsnIleHisGlnGluAspGlyMetGluGlnProMetGlnAs 17

581 ATGGCAATATTCACGAGAACGAGAGATGGAGCAGCCTATGCAGAA 532

17 nGlyGluAspArgProLeuGlyGlyGlyHisGlnProAlaG 34

531 TGGAGAGAAACCCCTTTGGGAGGAGTGAAGGCCACGAGCTGCAG 482

34 LyAsnArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTTPala 50

481 GAAATCGACGGGAGCCCTTCGCCACTTATTTTCATTCATGGGCC 432

51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPhe 67

431 ATACCCATAGGCAGATCAATGATGGGATGGTGGAGATGGAGATGAT 382

67 tGluIlePheMetGluGluMetArgGluIleArgGlyLeuArgGluL 84

381 GGAATATTCATGGAGGATGAGAGAAATCAGAGAAACTTAGGGAGC 332

84 euGlnLeuArgAsnCysLeuArgGlyLeuMetGlyGlyLeuSerAsnHis 100

331 TGCAGTTGAGGAATTTGTCGCTATCTTATGGGGAGCTCTCTAATC 282

101 HisAspHisHisAspGluPheCysLeuMetPro 111

281 CATGACCATCATGATGAATTTGCCCTTATGCCT 249

seq_name: gb_est1.A1193112

seq_documentation_block:

LOCUS A1193112 615 bp mRNA EST

DEFINITION qe69g09.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone

IMAGE:1744288 3' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0

KD PROTEIN HGR74 (HUMAN); mRNA sequence.

ACCESSION A1193112

VERSION A1193112.1 GI:3744321

KEYWORDS EST.

SOURCE Human

ORGANISM Homo sapiens

Eukaryotes: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;

Mammalia: Eutheria: Primates; Catarrhini, Homiidae, Homo.

1 (bases 1 to 615)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapb@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert length: 1.25 kb. Vector: pCMV-SPORT6

Size: 1.25 kb. Average insert size 1.25 kb. Life Technologies

catalog #: 11547-015"

High quality sequence stop: 445.

Location/Qualifiers

1. .615

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1744288"

/clone_lib="Soares_fetal_lung_NbHL19W"

/dev_stage="19 weeks (colicin resistant)"

/vector="pCMV-SPORT6"

/site="NotI; Cloned unidirectionally. Primer:

modified polylinker; Site: 1: NotI; Site: 2: Eco RI; 1st

strand cDNA was primed with a NotI - oligo(dT) primer

[5'-TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTTTTTTTTT-3']

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with NotI and cloned into

the NotI and Eco RI sites of a modified pT73 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by Bento

Soares and M. Patricia Bonaldo. This library was constructed

from a cDNA library as the fetal heart library, Soares

fetal heart NbHL19W."

BASE COUNT 169 a 145 c 121 g 180 t

ORIGIN

alignment_scores:

Quality: 609.00 Length: 111

Ratio: 5.486 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x A1193112/rev ..

Align seg 1/1 to reverse of: A1193112 from: 1 to: 615

1 MetAlaAsnIleHisGlnGluAspGlyMetGluGlnProMetGlnAs 17

584 ATGGCAATATTCACGAGAACGAGAGATGGAGCAGCCTATGCAGAA 535

17 nGlyGluAspArgProLeuGlyGlyGlyHisGlnProAlaG 34

534 TGGAGAGAAACCCCTTTGGGAGGAGTGAAGGCCACGAGCTGCAG 485

34 LyAsnArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTTPala 50

484 GAAATCGACGGGAGGAGCTTCGCCACTTATTTTCATTCATGGGCC 435

51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPhe 67

434 ATACCCATAGGCAGATCAATGATGGGATGGTGGAGATGGAGATGAT 385

67 tGluIlePheMetGluGluMetArgGluIleArgGlyLeuArgGluL 84

384 GGAATATTCATGGAGGATGAGAGAAATCAGAGAAACTTAGGGAGC 335

84 euGlnLeuArgAsnCysLeuArgGlyLeuMetGlyGlyLeuSerAsnHis 100

334 TGCAGTTGAGGAATTTGTCGCTATCTTATGGGGAGCTCTCTAATC 285

101 HisAspHisHisAspGluPheCysLeuMetPro 111

enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (INCIIS-Area Science Park, Trieste, Italy). 2 others

BASE COUNT 178 a 143 c 169 g 170 t

ORIGIN

alignment_scores: Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-327-750D-13 x A1929106 ..

Align seg 1/1 to: A1929106 from: 1 to: 662

```
1 MetAtaAsn1leHsGlnGluAsnGluMetGluGlnProMetGlnAs 17
96 ATGGCAATATTCACAGCAAGAAACAGAGATGGAGCAGCTATGCAGAA 145
17 nGlyGluGluAspArgLeuGlyGlyGlyGlyHisGlnProAlaG 34
146 TGGAGAGAGACCGCCCTTTGGGAGGAGTGAGGCCACCCCTTCGAG 195
34 lYAsnATgATgGlnGlnAlaArgATgLeuAlaProAsnPheATgTpAla 50
196 GAAATGACGGGACAGGCTCGCCGACTTCCCTTAATTTTCGATGGCC 245
51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
246 ATACCAATAGGAGATCAATAGGGATGGTGGAGATGGAGATGATAT 295
67 tGluIlePheMetGluGluMetArgGluIleArgATgGlyAspGlyAspPme 84
296 GGAATATTCATGGAGGAGATGAGAGAAATCAGAGAAACTTAGGGAGC 345
84 euGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
346 TGCAGTTCAGGAAATGTCTGCGTATCCTTATGGGGAGCTCTTAATC 395
101 HIsAspHisAspGluPheCysLeuMetPro 111
396 CATGACCATCATGATGAATTTTCCTTATGCTT 428
```

seq_name: gb_est2:BG745809

seq_documentation_block:
LOCUS BG745809 674 bp mRNA EST 15-MAY-2001
DEFINITION 60773951T1 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:4850293 3', mRNA sequence.

ACCESSION BG745809
VERSION BG745809.1 GI:14056462
KEYWORDS EST
SOURCE human

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 674)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.
Eukaryotic Genome Project
Tissue Procurement: Dr. Mack Watson
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM1691 row: e column: 14
High quality sequence stop: 674.

FEATURES

source
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4850293"
/lab_host="NIH_MGC_113"
/note="Organ: spleen; Vector: pOTB7; site_1: XhoI; site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACAG(G). Library constructed by Ling Hong in the Laboratory of Genet. Molec. Biol. (Rubin Laboratory), University of California, San Diego, CA 92093. Cloning strategy: Stratagene SuperScript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 167 a 186 c 142 g 179 t

alignment_scores: Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-327-750D-13 x BG745809/rev ..

Align seg 1/1 to reverse of: BG745809 from: 1 to: 674

```
1 MetAtaAsn1leHsGlnGluAsnGluMetGluGlnProMetGlnAs 17
529 ATGGCAATATTCACAGCAAGAAACAGAGATGGAGCAGCTATGCAGAA 480
17 nGlyGluGluAspArgLeuGlyGlyGlyGlyHisGlnProAlaG 34
479 TGCAGAGAGACCGCCCTTTGGGAGGAGTGAGGCCACCCCTTCGAG 430
34 lYAsnATgATgGlnGlnAlaArgATgLeuAlaProAsnPheATgTpAla 50
429 GAAATGACGGGACAGGCTCGCCGACTTCCCTTAATTTTCGATGGCC 380
51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
379 ATACCAATAGGAGATCAATAGGGATGGTGGAGATGGAGATGATAT 330
67 tGluIlePheMetGluGluMetArgGluIleArgATgGlyAspGlyAspPme 84
329 GGAATATTCATGGAGGAGATGAGAGAAATCAGAGAAACTTAGGGAGC 280
84 euGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
279 TGCAGTTCAGGAAATGTCTGCGTATCCTTATGGGGAGCTCTTAATC 230
101 HIsAspHisAspGluPheCysLeuMetPro 111
229 CATGACCATCATGATGAATTTTCCTTATGCTT 197
```

seq_name: gb_est1:AV701506

seq_documentation_block:
LOCUS AV701506 720 bp mRNA EST 08-OCT-2000
DEFINITION AV701506 ADS Homo sapiens cDNA clone ADBBH12 5', mRNA sequence.

ACCESSION AV701506
VERSION AV701506.1 GI:10717836
KEYWORDS EST
SOURCE human

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 720)
PENG, Y., SONG, H., HUANG, Q., GU, Y., YANG, Y., GAO, G., XIAO, H., XU, X., LI, N., QIAN, B., LI, J., QIAN, X., CHENG, Z., XU, Z., ZENG, L., XU, S., GU, W., TU, Y., JIA, J., FU, G., REN, S., ZHONG, M., LU

```

84  eugluLeuArgAncCysLeuArgIleLeuMetGlyGluLeuSerAanHis 100
336  TCGAGTTGAGGAATTCGTCTGCTATCTTATGGGAGGCTCTCTAATCAC 287
101  HisAspHisHisAspGluPheCysLeuMetPro 111
286  CATGCCATCATCATGAATTTGGCTTATGCGCT 254
seq_name: gb_est2:BG482054

seq_documentation_block:
LOCUS      BG482054          557 bp    mRNA           21-MAR-2001
DEFINITION 602527934F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4651353 5',
mRNA sequence.
ACCESSION  BG482054
VERSION    BG482054.1 GI:13414333
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  NIH-MGC http://mgi.nci.nih.gov/
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            Cloned by: Incyte Genomics, Inc.
            Clon strategy: Lambda phage. Clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LCM1434 row: d column: 10
            High quality sequence stop: 556.
            Location/Qualifiers
                1..557
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="NIH_MGC_21"
                /lab_host="DH10B"
                /note="Organ: placenta; Vector: pBluescript"
                Site 2: EcoRI; cDNA made by oligo-dT priming
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Size-selected for
                for average insert size 1.8kb. Library constructed by
                Ling Hong in the laboratory of Gerald M. Rubin (University
                of California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT  157 a 126 c 166 g 108 t
ORIGIN
alignment_scores:
Quality: 597.00 Length: 111
Ratio: 5.427 Gaps: 0
Percent Similarity: 99.099 Percent identity: 99.099

alignment_block:
US-09-327-750D-13 x BG482054
Align seg 1/1 to: BG482054 from: 1 to: 557
1 MetAlaAsnIleHisGluGluLeuMetGluGluMetGluGluProMetGlnAs 17
185 ATGGCAATATTCACGAGAAACCAAGAGATGGGACGCTATGCAGAA 234
17 nGlyGluAspArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 34
235 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284

```

```

34  lvaAsnArgArgGlyGluAlaArgArgGluLeuAlaProbenPheArgTgPal 50
285 GAATTCGAGCGGAGCGCTCGCGGACTTGCCTTAATTTTCGATCGGCC 334
51  tleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspMe 67
335 ATACCAATAGGACAGATCAATGATGGATGGGTGGAGATGAGATGATAT 384
67  GtullePheMetGluGluMetArgGluIleArgArgGlyLysLeuArgGluL 84
385 GGAATATTCATGGAGAGATGAGAGAAATCAGAGAAACTTAGGGAGC 434
84  eugluLeuArgAncCysLeuArgIleLeuMetGlyGluLeuSerAanHis 100
435 TGCAGTTGAGGAATTCGTCTGCTATCTTATGGGAGGCTCTCTAATCAC 484
101  HisAspHisHisAspGluPheCysLeuMetPro 111
485 CATGCCATCATCATGAATTTGGCTTATGCGCT 517
seq_name: gb_est2:BG705685

seq_documentation_block:
LOCUS      BG705685          567 bp    mRNA           07-MAY-2001
DEFINITION 602668910P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4791506 5',
mRNA sequence.
ACCESSION  BG705685
VERSION    BG705685.1 GI:13980274
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  NIH-MGC http://mgi.nci.nih.gov/
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NIGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LCM1434 row: d column: 03
            High quality sequence stop: 567.
            Location/Qualifiers
                1..567
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="NIH_MGC_96"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pBluescriptR (modified
                pBluescript RSV); Site 1: BamHI; Site 2: SalI-XhoI (gtcggg
                gtcggg) primer 5' primer 3' primer 5' primer 3'
                size-selected for average insert size 2.3 kb and
                normalized to 500 ng. This is a primary library
                for full-length clones and constructed using the
                Cap-trapper method (Carninci, in preparation). Library
                constructed by M. Brownstein (NIGRI/NIGRI, National
                Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT  160 a 130 c 168 g 109 t
ORIGIN
alignment_scores:
Quality: 597.00 Length: 111

```

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This sequence can be seen in the following URL
 Project URL: <http://ludwig.org.br/scripts/gethtml2.pl?tl=PM0&tl2=PM0-HT0911-080201-012-hl043-2001-02-08&f4=1>
 Seq primer: puc 18 forward
 High quality sequence stop: 550.
 Location/Qualifiers
 1..550
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_16"
 /site="organ: head,neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions.
 BASE COUNT 154 a 113 c 137 g 146 t
 ORIGIN
 alignment_scores:
 Quality: 596.00 Length: 112
 Ratio: 5.369 Gaps: 1
 Percent Similarity: 99.107 Percent Identity: 99.107
 alignment_block:
 US-09-327-750D-13 x BG993540 ..
 Align seg 1/1 to: BG993540 from: 1 to: 550
 1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
 48 ATGCCAATATTACCCAGGAACCAAGAGATGGCAGCCTATGCAGAA 97
 17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyGlyHisGlnProAlaG 34
 98 TGGAGAGGAAGACCGCCCTTTGGGAGAGGTGAAGCCACCAGCCTGCAG 147
 34 lyAsnArgArgGlyGlnAlaArgLeuAla.ProAsnPheArgTrpAla 50
 148 GAATTCGACGGGACAGGCTGGCCGCTTCGCCCTTAATTTTCGATGGGC 197
 50 alleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspM 67
 198 CATACCAATAGCCAGACATCAATGATGGGATGGGTGGAGATGGATGATA 247
 67 etGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGlu 83
 248 TGGAAATATTATCGAGGAGATGAGAGAAATCAGAAATCAGAAATCAGGAG 297
 84 LeuGlnLeuArgPheCysLeuArgGluLeuMetGlyGluLeuSerAsnH 100
 298 CTGCAATTGAGCAATTGTCTGCGTATCTCTATGGGAGCTCTCTAATCA 347
 100 sHisAspHisHisAspClyPheCysLeuMetPro 111
 348 CCATGACCATCATGATGAATTTTGGCTTATGCTT 381
 seq_name: 9b_est1:BE252272
 seq_documentation_block:
 BE252272 559 bp mRNA EST 13-JUL-2000
 DEFINITION 60114115F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3354819 5',
 mRNA sequence.
 ACCESSION BE252272
 VERSION BE252272.1 GI:912409
 KEYWORDS EST.

SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 559)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@nih.gov
 Tissue: prostate; ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCMI57 row: n column: 04
 High quality sequence stop: 559.
 Location/Qualifiers
 1..550
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_16"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pORF7; Site_1: XbaI; Site_2:
 EcoRI; CDNA made by oligo-dr primers. Directionally
 cloned. Incorporated into the following s,
 and the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using Zap-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 157 a 126 c 166 g 110 t
 ORIGIN
 alignment_scores:
 Quality: 591.00 Length: 111
 Ratio: 5.429 Gaps: 0
 Percent Similarity: 98.198 Percent Identity: 98.198
 alignment_block:
 US-09-327-750D-13 x BE252272 ..
 Align seg 1/1 to: BE252272 from: 1 to: 559
 1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
 177 ATGGCAATATTACCCAGGAACCAAGAGATGGCAGCCTATGCAGAA 226
 17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGlnProAlaG 34
 227 TGGAGAGGAAGACCGCCCTTTGGGAGAGGTGAAGCCACCAGCCTGCAG 276
 34 lyAsnArgArgGlyGlnAlaArgLeuAla.ProAsnPheArgTrpAla 50
 277 GAATTCGACGGGACAGGCTGGCCGCTTCGCCCTTAATTTTCGATGGGC 326
 51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspM 67
 327 ATACCAATAGCCAGACATCAATGATGGGATGGGTGGAGATGGATGATA 376
 67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGlu 84
 377 GGAATATTATCGAGGAGATGAGAGAAATCAGAAATCAGAAATCAGGAGC 426
 84 euGlnLeuArgAsnCysLeuArgLysLeuMetGlyGluLeuSerAsnHis 100
 427 TGCAGTTGAGGAATGCTTCGCTTCCTTATGGGAGGCTCTCTATCAGC 476
 101 HisAspHisHisAspClyPheCysLeuMetPro 111

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 11, 2002, 14:17:05 ; Search time 1319.57 Seconds
(without alignments)
8751.355 Million cell updates/sec

Title: US-09-327-750D-28
Perfect score: 700
Sequence: 1 acagagctgctgcagcagc.....ggggcttgctgcagtgga 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenBank.*

1: gb-ba.*

2: gb-hg.*

3: gb-in.*

4: gb-om.*

5: gb-pat.*

6: gb-pl.*

7: gb-pr.*

8: gb-ro.*

9: gb-sy.*

10: gb-un.*

11: gb-vi.*

12: gb-wt.*

13: gb-zo.*

14: gb-ec.*

15: gb-ec.*

16: gb-ec.*

17: gb-ec.*

18: gb-ec.*

19: gb-ec.*

20: gb-ec.*

21: gb-ec.*

22: gb-ec.*

23: gb-ec.*

24: gb-ec.*

25: gb-ec.*

26: gb-ec.*

27: gb-ec.*

28: gb-ec.*

29: gb-ec.*

30: gb-ec.*

31: gb-ec.*

32: gb-ec.*

33: gb-ec.*

34: gb-ec.*

35: gb-ec.*

36: gb-ec.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	692	98.9	700	10	AF187066	AF187066 Mus muscu
2	676	96.6	854	10	AF097440	AF097440 Mus muscu
3	355.8	50.8	519	10	AF187065	AF187065 Rattus no
4	337.8	48.3	793	9	BC003190	BC003190 Homo sapi
5	307.4	43.9	891	9	AF187068	AF187068 Homo sapi
6	307.4	43.9	891	9	AF187068	AF187068 Homo sapi
7	307.4	43.9	891	9	AF187068	AF187068 Homo sapi
8	154.8	22.1	785	10	AF097439	G22708 MARC 4953-4
9	154.8	22.1	785	10	AF097439	AF097439 Mus muscu
10	101.4	14.5	477	11	G35294	G35294 Human DNA s
11	101.4	14.5	477	11	G35294	G35294 Human DNA s
12	80.2	11.5	1229	9	AK000959	AK000959 Homo sapi
13	80.2	11.5	1229	9	AK000959	AK000959 Homo sapi
14	80.2	11.5	1229	9	AK000959	AK000959 Homo sapi
15	77.4	11.1	244	11	GI3350	GI3350 Human STRS
16	57.6	9.7	799	6	AF097438	AF097438 Mus muscu
17	57.6	9.7	799	6	AF097438	AF097438 Mus muscu
18	57.6	9.7	799	6	AF097438	AF097438 Mus muscu
19	58	8.3	828	9	AF183416	AF183416 Homo sapi
20	58	8.3	828	9	AF183416	AF183416 Homo sapi
21	57.2	8.2	43952	9	HS10564	AF220189 Homo sapi
22	56.4	8.1	791	9	AF237783	AF237783 Homo sapi
23	55.8	8.0	744	9	AF251053	AF251053 Homo sapi
24	55.8	8.0	744	9	AF251053	AF251053 Homo sapi
25	55.8	8.0	744	9	AF251053	AF251053 Homo sapi
26	55.8	8.0	744	9	AF251053	AF251053 Homo sapi
27	54.4	7.9	40584	1	AF133348	AF133348 Human DNA
28	54.4	7.9	40584	1	AF133348	AF133348 Human DNA
29	54.4	7.9	40584	1	AF133348	AF133348 Human DNA
30	54	7.7	838	10	AF051347	AF051347 Mus muscu
31	54	7.7	838	10	AF051347	AF051347 Mus muscu
32	49.2	7.0	18306	2	AC013570	AC013570 Homo sapi
33	48.8	7.0	37224	3	AC007061	AC007061 Leishmani
34	48.2	6.9	36823	9	HSU20961	HSU20961 Human DNA s
35	46	6.6	154577	2	AL357046	AL357046 Homo sapi
36	46	6.6	154577	2	AL357046	AL357046 Homo sapi
37	46	6.6	154577	2	AL357046	AL357046 Homo sapi
38	44.8	6.4	18257	9	AL353508	AL353508 Homo sapi
39	44.8	6.4	18257	9	AL353508	AL353508 Homo sapi
40	44.2	6.3	172000	2	LMV23815_5	Continuation (6 of
41	43.2	6.2	9211	3	AB030033	AB030033 Dictyoste
42	43.2	6.2	210321	2	AC079438	AC079438 Mus muscu
43	42.8	6.1	1580	3	PFACSPR	PFACSPR Plasmodium
44	42.6	5.1	164025	2	AC022562	AC022562 Homo sapi
45	42.6	5.1	166922	2	ALS12288	ALS12288 Homo sapi

ALIGNMENTS

RESULT	1
LOCUS	AF187066 700 bp mRNA 11-JUN-2000
DEFINITION	Mus musculus p75NTR-associated cell death executor (Nade) mRNA, complete cds.
ACCESSION	AF187066
VERSION	AF187066.1 GI:8452897
KEYWORDS	Complete cds.
ORGANISM	Mus musculus
REFERENCE	1 (bases 1 to 700)
AUTHORS	Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D., Suvento,P., Hanaka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.
TITLE	Nade, a p75NTR-associated cell death executor, is involved in signal transduction mediated by the common neurotrophin receptor
JOURNAL	J Biol Chem. 275 (23), 17566-17570 (2000)
HEADLINE	2028829

[illegible]

where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Embl: EMBL; Svi: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at <http://www.sanger.ac.uk/WormPEP/>

Genbank: Genbank; J: J-Express; Tr: TrEMBL; Wp: WormPEP; Information from the human BAC library described in U-J. Kim et al. (1996)

Genomics 34, 213-218.

VECTOR: phelocAC11

This sequence is the entire insert of clone CTA-714B7.

Location/Qualifiers

1. .98274

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="12"

/map="CTA-714B7"

/clone_id="CT978SK-A2"

109..136

region

183..258

region

note="14 copies 2 mer ac 100 conserved"

183..258

region

note="LMCS repeat: matches 7840. .7913 of consensus"

1177..1446

nature

/note="match: GSS: Em:AQ373724"

1301..1336

region

note="1336 repeat: matches 2693. .2750 of consensus"

2458..5038

region

note="LMND1 repeat: matches 3673. .6170 of consensus"

5036..5486

region

note="LMEC repeat: matches 2052. .2207 of consensus"

5436..6885

region

note="LMND repeat: matches -1. .1538 of consensus"

6896..7185

region

note="L2 repeat: matches 2356. .2710 of consensus"

complement(7187..7611)

nature

7189..7594

region

note="MIR repeat: matches 3600. .256 of consensus"

complement(7713..8092)

nature

/note="match: GSS: Em:AQ071449"

7904..7978

region

note="MIR repeat: matches 20. .95 of consensus"

8086..8300

region

note="MIR repeat: matches 8. .255 of consensus"

8134..8565

nature

8519..8828

region

note="ALUSx repeat: matches 1. .280 of consensus"

complement(8897..9290)

nature

/note="match: GSS: Em:AQ343214"

9235..9361

region

note="L2 repeat: matches 2619. .2749 of consensus"

9378..9890

nature

/note="match: GSS: Em:AQ284832"

10135..10428

region

note="L2 repeat: matches 1. .294 of consensus"

10429..10563

region

note="ALUS/x repeat: matches 1. .135 of consensus"

10811..10868

region

note="L2 repeat: matches 2648. .2705 of consensus"

10819..11064

region

note="MIR repeat: matches 2. .255 of consensus"

11112..11324

region

note="MIR repeat: matches 11. .255 of consensus"

11869..12012

region

note="L2 repeat: matches 20. .167 of consensus"

12925..13223

region

note="ALU repeat: matches 1. .300 of consensus"

13290..13495

region

Tel: 4157259687
 Fax: 4157259689
 Email: myers@hgc.stanford.edu

Primer A: AACATCTTCCATGAAGTTGATG
 Primer B: CTTTGGCATCTTCTGCAA
 STS size: 106
 PCR Profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 15 seconds
 Annealing: 60 degrees C for 30 seconds
 Polymerization: 72 degrees C for 23 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600

Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Amplifrag Gold Polymerase: 0.07 units/uL
 Total Vol: 5 uL

Buffer: MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from NS1315
 -- Washington University/Merck EST sequence.

FEATURES

source location/qualifiers
 STS
 primer_bind 44..149
 primer_bind complement(130..149)
 BASE COUNT 153 a 115 c 79 g 130 t
 ORIGIN
 Query Match 14.5%; Score 101.4; DB 11; Length 477;
 Best Local Similarity 76.3%; Prev. No 4.5e-16;
 Matches 161; Conservative 0; Mismatches 51; Indels 9; Gaps 2;
 Oy 339 cgacctgcccttaactccgatggccattcccaacagcgc-----agataaatagcggg 392
 Db 438 CCACCTCCCGCGGTTTGGAGAGGATGCGCCCAATAGCTGTGCGATAACATTGTATG 379
 Oy 393 ttgggtggagatgatgataatgaaatgctatggagagatgagagatccggaga 452
 Db 378 ATAGATGGAGATGGAGATGATGACCGGTTCATGGAGGAGATGAGAGCTTAGGAGG 319
 Oy 453 aagcttagagagatcagctcagaaattgtctcgcacctctatggggagcgtctaac 512
 Db 318 AAAATTAGGGAACCTCAGTTGAGGTACAGTCTGCGATCTTATAGGGGACC---CTCCT 262
 Oy 513 caccacagaccatgatgatctctgccttattccttgaact 553
 Db 261 CACCATGATCATCATGATGATGTTTGCCTTATGCTTGAAT 221

RESULT 11

HSV351F8
 LOCUS HSV351F8 45678 bp DNA PRI 23-NOV-1999
 DEFINITION Human DNA sequence from cosmid V351F8, between markers DX3366 and DX587 on chromosome X contains ESTs.
 ACCESSION 270719
 VERSION 270719.1 GI:1261915
 KEYWORDS X.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 45678)
 AUTHORS Whiteley M.
 TITLE Direct Submission
 JOURNAL Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, UK. E-mail enquiries: humquerry@sanger.ac.uk
 COMMENT IMPORTANT: This sequence is the entire insert of clone V351F8. The true left end of clone V351F8 is at 1 in this sequence. The true right end of clone V351F8 is at 45678.
 V351F8 is from the human chromosome X-specific cosmid library.

FEATURES
 source location/qualifiers
 1..45678 "Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="X"
 /clone="GHC-351F8"
 /clone_lib="SCcv"
 1..70
 /partial
 repeat_region
 /note="Alu repeat: matches 80..11 of consensus"
 642..930
 repeat_region
 /note="Alu repeat: matches 1..308 of consensus"
 7406..12179 element fragment
 5243..5276
 repeat_region
 /note="17 copies of 2 mer 82 & conserved"
 5279..5567
 repeat_region
 /partial
 /note="Alu repeat: matches 304..11 of consensus"
 9975..10369
 repeat_region
 /note="MSTA element fragment"
 10399..12017
 repeat_region
 /note="MSTAR element fragment"
 1318..13578
 repeat_region
 /note="THE18 element fragment"
 13715..11846 element fragment
 /note="MT18 element fragment"
 12074..12170
 repeat_region
 /note="MSTA element fragment"
 12227..12486
 repeat_region
 /note="MSTA element fragment"
 12437..12502
 repeat_region
 /note="MSTC element fragment"
 12756..13059
 repeat_region
 /partial
 /note="Alu repeat: matches 308..11 of consensus"
 14478..14650
 repeat_region
 /note="THE18 element fragment"
 14491..14536
 repeat_region
 /note="MSTA element fragment"
 14715..14831
 repeat_region
 /note="THE18 element fragment"
 15099..15533
 repeat_region
 /note="L1 element fragment"
 15534..15626
 repeat_region
 /note="31 copies of 3 mer 85 & conserved"
 15566..15626
 repeat_region
 /note="3 copies of 18 mer 98 & conserved"
 16332..16436
 repeat_region
 /note="L1 element fragment"
 16728..16763
 repeat_region
 /note="9 copies of 4 mer 94 & conserved"
 16764..17054
 repeat_region
 /partial
 /note="Alu repeat: matches 308..11 of consensus"
 19420..19449
 repeat_region
 /note="15 copies of 2 mer 87 & conserved"
 complement(19537..20013)
 misc_feature
 /note="match 3' EST NS1315 clone 283089"
 20071..20114
 repeat_region
 /note="22 copies of 2 mer 98 & conserved"
 20073..20112
 repeat_region

```
Db 256 AGGAGGGAGCAGGCCGCCACGACAGATGAAGAAGATCCCGCCATTTGGGAGGGGGTGA 315
Qy 260 gggccaccgctgctgctgcaacaacacacacacacacacacacacacacacacacacac 319
Db 316 AGCCAGAGCCTGGAGGAATATCA----- 341
Qy 320 ccgaagagggccaggtgcgcagactgcgccttaactccctgcagtgccgacattcccaacagga 379
Db 342 --GCGGGGGCGAGTTAGCGGACCTTTCCTTAATTTTCGATGGCCCACTCATATAGGCA 399
Qy 380 gatgatcgaggggtgggtggagatgagatggaatgcttctcagtgagagatgag 439
Db 400 TATTGAGCACAA-----TGAAGCGAGAGATGATGAGAAAGTTTGTAGGCGAGATAT 453
Qy 440 agagatcccgagagactgagagatgacagtcagagaaattgtctacgacatccttatggg 499
Db 454 GGAATCAAGAGAAAGACTAGGGAACAGACGATGAGCGACTATATGCGCTTCCAAA---- 509
Qy 500 ggaagctgtctaacacccagatcacacacagatgagtaattgccttatgccttaactcgttc 559
Db 510 -----CTCTGACACCTGACACACCTTTCCTTAATTTTCGATGAGTTCGATGATCTTAA 561
Qy 560 attcccccctgacatcagtcagtcgcctgcgtgcagcccttccctcctcctccttcttc 618
Db 562 AGTTTTCGCTGAGGTAAATGACACACTGCTTTACAAAGCTTGATTTTGTGATTACT 620

RESULT 13
AX100231 1364 bp DNA PAT 02-APR-2001
SOURCE Homo sapiens
ORGANISM human
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tang,Y.T., Hillman,J.L., Yue,H., Reddy,R., Lal,P., Shah,P.,
Azimzai,T., Baughn,M.R., Lo,D.A., Bandman,O., Shih,L.B. and
Peterson,G. Associated with cell differentiation
TITLE Patent: WO 0119860-A 32 22-MAR-2001;
JOURNAL Incyte Genomics, Inc. (US)
FEATURES
    source          1..1364
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
BASE COUNT 411 a 269 c 322 g 362 t
ORIGIN
Query Match 11.5%; Score 80.2; DB 6; Length 1364;
Best Local Similarity 56.2%; Pred. No. 1; Re-10;
Matches 269; Conservative 0; Mismatches 153; Indels 57; Gaps 4;
Qy 143 gccagagatgacacacaggaacaaatctcatggtcccaatgtccacaggaacagca 202
Db 255 GTCCAAAGAGGACTAGCGCAACAAATCTCAACGGGGAATGCCCAACAGAAAGCA 314
Qy 203 agagctggagcag---ccctcgagatggcaggaacacccctgtggtggagaggtga 259
Db 315 AGAGGGGAGCGGCCCCACGACGATGAGAGATATCCGCCATTTGGAGGGGGTGA 374
Qy 260 gggccaccgctgctgcaacaacacacacacacacacacacacacacacacacacac 319
Db 375 AGGCCAGAGCCTGGAGGAATATCA----- 400
Qy 320 ccgaagagggccaggtgcgcagactgcgccttaactccctgcagtgccgacattcccaacagga 379
Db 401 --GCGGGGGCGAGTTAGCGGACCTTTCCTTAATTTTCGATGGCCCACTCATATAGGCA 458
```

```
Qy 380 gatgatcgaggggtgggtggagatgagatgataatgaaattgtctacgacatccttatggg 439
Db 459 TATTGAGCACAA-----TGAAGCGAGAGATGATGATGAGAAAGTTTGTAGGCGAGATAT 512
Qy 440 agagatcccgagagagctgagagctacagtcagagaaattgtctacgacatccttatggg 499
Db 513 GGAATCAAGAGAAAGACTAGGGAACAGACGATGAGCGACTATATGCGCTTCCAAA---- 568
Qy 500 ggaagctgtctaacacccagatcacacacagatgagtaattgccttatgccttaactcgttc 559
Db 569 -----CTCTGACACCTGACACCACTTATGACTTTTGCTCATACCTTGATCTCTAAA 620
Qy 560 attcccccctgacatcagtcagtcgcctgcgtgcagcccttccctcctcctccttcttc 618
Db 621 AGTTTTCGCTGAGGTAAATGACACACTGCTTTACAAAGCTTGATTTTGTGATTACT 679

RESULT 14
H5353G.9 69648 bp DNA PBI 23-NOV-1999
SOURCE Homo sapiens
ORGANISM human
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 69648)
Direct Submission
TITLE Bird C.
JOURNAL Submitted (02-JUN-1999) Sanger Centre, Hinxton, Cambridgeshire,
CS10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clones@sanger.ac.uk
On May 11, 1999 this sequence version replaced gi:4678749.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations. The
differences are then compared to the reference sequence. If the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
En: EMBL; Sw: SWISSPROT; Tr: TrEMBL
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unseq'
key.
This sequence was generated from part of bacterial clone contigs of
human chromosome X, constructed by the Sanger Centre Chromosome X
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
635G19 is from the library RPC14 constructed at the Roswell Park
Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/VECTOR:pcYPAC2
IMPORTANT: This sequence is not the entire insert of clone 635G19.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring subclones.
This is the end of clone U65M4 (Z81014) is at 69549 in this
sequence. The true right end of clone U10D3 (Z85997) is at 100 in
this sequence.
FEATURES
    source          1..69648
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
```

```

PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
  Template: 25 ng
  Primer: each 1 uM
  dNTPs: each 200 uM
  Taq Polymerase: 0.05 units/ul
  Total Vol: 10 ul

  Buffer:
    MgCl2: 2.5 mM
    KCl: 50 mM
    Tris-HCl: 20 mM
    pH: 8.3

  Prepared with primer pairs derived from M38188 -- Unigene.

FEATURES
  source 1..244
    STS
    primer_bind 84..103
    primer_bind 84..103
    primer_bind complement(199..221)
  BASE COUNT 57 a 50 c 42 g 95 t
  ORIGIN

Query Match 11.18; Score 77.4; DB 11; Length 244;
Best Local Similarity 80.24; Freq. No. 8.6e-10;
Matches 105; Conservative 0; Mismatches 21; Indels 5; Gaps 1;

QY 567 ctgagctcaactgactcgcctccgccttcctccgctttctcgcagctacc 626
Ddb 15 CATGAGATTAATACTGATGATCCCGCTTTCTTTTCTCTGCTATTTTCCATATGCG 74

QY 627 tttaatgaccgcttggtagcgcctggtttattccatgcgcattgcgcaggggggct 686
Ddb 75 TTACTGATCCGTTCTGCTGTAACCCATGTTTATT-----CCATGTGTCAAGTGGGTCT 129

QY 687 tgtgtgccag 697
Ddb 130 TGTG66GGCAG 140

Search completed: March 11, 2002, 16:33:03
Job time: 8:58 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 13:08:25 ; Search time 1609.39 Seconds
(without alignments)
4673.848 Million cell updates/sec

Title: US-09-327-750d-28
Perfect score: 700
Sequence: 1 agggagctgtgcccagcagc.....ggggctgtgtgcccagcaga 700

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST.*

- 1: em_estfun.*
- 2: em_esthum.*
- 3: em_esthum.*
- 4: em_esthum.*
- 5: em_estpl.*
- 6: em_estba.*
- 7: em_estro.*
- 8: em_estov.*
- 9: em_hic.*
- 10: gb_estl.*
- 11: gb_estl.*
- 12: gb_estl.*
- 13: gb_gss.*
- 14: gb_gss.*
- 15: em_gss_hum.*
- 16: em_gss_inv.*
- 17: em_gss_pln.*
- 18: em_gss_pro.*
- 19: em_gss_rod.*
- 20: em_gss_vrt.*
- 21: em_gss_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	687.4	98.2	809	12	AK004531	Mus muscu
2	687.4	98.2	845	12	AK010500	Mus muscu
3	687.7	91.0	787	11	BF608455	Wt1_00139
4	627.6	89.7	642	10	A1118980	ue93c08.y
5	606.8	86.7	628	10	AM476468	uq76d02.y
6	591.8	84.4	599	11	BC088461	AA272375
7	590.4	84.3	612	10	AA272375	VB52907.r
8	585.4	83.6	826	12	AK003294	Mus muscu
9	577.8	82.5	817	11	BC107854	604010.y
10	571.6	82.5	817	11	BC107854	604010.y
11	566.6	80.9	923	10	BF178306	501808545
12	535.6	76.5	542	10	BE334866	us90all.y

13	533.2	76.2	810	11	BF168928	601775392
14	533.2	76.2	918	11	BF1249717	602996190
15	531.4	75.9	557	10	AA215070	mu76f04.f
16	524.8	75.0	577	10	AA268306	va89c10.f
17	523.6	74.8	538	10	A1006575	ue13606.y
18	516	73.7	590	11	W647111	mu76f04.f
19	510.4	72.9	347	10	AA050176	mu14b06.y
20	509.8	72.7	577	11	BF681983	602116954
c	507	72.4	536	11	W88871	mf89d04.r1
22	507	72.4	536	11	W81757	me95d06.r1
23	507	72.4	536	11	W81757	me95d06.r1
24	497.4	71.1	542	10	AA004191	mg82d06.f
25	496.8	71.0	1042	11	W15845	mb5b09.r1
26	492.4	70.3	503	10	AW08751	uf57a05.y
27	489.2	69.9	853	11	BF082462	602877892
28	487	69.6	522	10	AA068855	mm64c09.f
29	485.4	69.3	506	10	AA019613	kl38b06.f
30	485.4	69.3	603	10	AW257617	kl38b06.f
31	483.8	69.1	533	10	AA050176	mu14b06.y
32	480	68.6	532	10	AA050176	mu14b06.y
33	479.8	68.5	534	10	AA214909	vm82d02.r1
34	479.8	68.5	534	10	AA259353	vm82d02.r1
35	476.6	68.1	533	11	W29675	mc07d02.r1
36	468.4	66.9	489	10	AA044548	ml12c02.f
37	466.6	66.7	560	10	AA870604	vq23d08.f
38	465	66.4	486	10	AA259786	vq23d08.f
39	461.2	65.9	509	10	AA636864	v120904.f
40	458	65.4	540	10	AA636864	v120904.f
41	453.4	64.8	562	11	BF681719	602116954
42	452.3	64.6	457	10	BE554418	u7-m-AH1-
43	452.3	64.6	457	10	BE554418	u7-m-AH1-
44	451	64.4	459	10	BE198421	u978d07.y
45	450.6	64.4	533	10	AA214902	mu76d05.f

ALIGNMENTS

RESULT 1
AK004531 809 bp mRNA 05-JUL-2001
DEFINITION Mus musculus 18 days embryo cDNA, RIKEN full-length enriched library, clone:1190011J23, full insert sequence.
ACCESSION AK004531
VERSION AK004531.1 GI:12835760
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA, clone:1190011J23.
ORGANISM Mus musculus
Mammalia; Eutheria; Chordata; Vertebrata; Euteleostomi; Rodentia; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 809)
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)
MEDLINE 92279253
PUBMED 10349636
REFERENCE 3 (bases 1 to 809)
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome research. 10 (10), 1617-1630 (2000)
PUBMED 11042159
AUTHORS Shibata, K., Itoh, M., Mizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Schmitt, A., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., Hayashizaki, Y., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Tokawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

554	TCGGTCATTCCGCCCTCAGATCCATCATCTGAGCTCCGCTGAGCCCTTCCTCGGCTT	613
Qy		
553	TCGGGTCAATCCGCCCTCCGAGATCCATCTGTGACTCCGCGTGTAGCCCTTTTCTCGCAT	612
Db		
RESULT 8		
AK001294		
LOCUS	826 bp	MRNA
DEFINITION	Mus musculus 18 days embryo cDNA, RIKEN full-length enriched library, clone 1110002F04 full insert sequence.	AK001294
		05-JUL-2001

AK003294	GI:12833872
ACCESSION	AK003294.1
VERSION	
KEYWORDS	CAP trapper.
SOURCE	Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA, clone:LIBRIKEN full-length enriched mouse cDNA library clone:ll002f04.
ORGANISM	Mus musculus
REFERENCE	Eukaryotic Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 826)
TITLE	Cardinci, P. and Hayashizaki, Y.
JOURNAL	High-efficiency full-length cDNA cloning
MEDLINE	Methods in enzymology. 303, 19-44 (1999)
PUBMED	99279253
REFERENCE	2 (bases 1 to 826)
AUTHORS	Cardinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, N., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to

TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome research. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
REFERENCE	10 (10), 1617-1630 (2000)
AUTHORS	Shibata,K., Itoh M., Aizawa, K., Nagaoke, S., Sasaki, N., Carninci, P., Kono, H., Ishiyama, J., Nishi, K., Kizunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishiyama, J., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome research. 10 (11), 1757-1771 (2000)
MEDLINE	10536821
REFERENCE	10 (11), 1757-1771 (2000)
AUTHORS	4 (1688as 1 to 826) The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 695-690 (2001)
REFERENCE	5 (bases 1 to 826)
AUTHORS	Adachi, J., Aizawa, K., Akehira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Haru, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Horii, F., Imoto, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koyas, S., Kurahara, C., Matsuyama, T., Okajima, Y., Oishi, K., Nomura, A., Numakura, A., Ono, M., Okazaki, Y., Oshida, K., Oshikawa, T., Shibata, K., Shibata, Y., Shiozaki, T., Sasaki, D., Suzuki, K., Shibata, M., Tagami, M., Takahashi, F., Sogabe, Y., Shiraki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyé, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: 41.45.503-9216.gsc.riken.go.jp/, Tel:81-45-953-9222, Fax:81-45-503-9216)
REFERENCE	Please visit our web site (http://genome.gsc.riken.go.jp/) for

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 11, 2002, 16:33:03 ; Search time 1319.57 Seconds
(without alignments)
11139.225 Million cell updates/sec

Title: US-09-327-750d-29
Perfect score: 891
Sequence: 1 acccaaccacccactctat.....ataaagcaatttaaaagc 891

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs. 824859755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenBank:

1: gb-ba.*

2: gb-htg.*

3: gb-in.*

4: gb-om.*

5: gb-ov.*

6: gb-ow.*

7: gb-ph.*

8: gb-pl.*

9: gb-pr.*

10: gb-ro.*

11: gb-sts.*

12: gb-sy.*

13: gb-un.*

14: gb-vi.*

15: en-ba.*

16: en-fun.*

17: en-lun.*

18: en-lun.*

19: en-om.*

20: en-or.*

21: en-ov.*

22: en-pat.*

23: en-ph.*

24: en-pl.*

25: en-ro.*

26: en-sts.*

27: en-sy.*

28: en-un.*

29: en-vi.*

30: en-htgo-hum.*

31: en-htgo-inv.*

32: en-htgo-rod.*

33: en-htg-hum.*

34: en-htg-inv.*

35: en-htg-rod.*

36: en-htg-other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the best result found, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	889.4	99.8	891	9	HUMOGC	M38188 Human unkno
2	889.4	99.8	891	9	AF187064	AF187064 Homo sapi
3	609.4	68.4	793	9	BC003190	BC003190 Homo sapi
4	363	40.7	854	10	AF097440	AF097440 Mus muscu
c	354.2	39.8	421	11	G72708	G72708 MARC 4953-4
5	312.2	35.0	200	10	AF187066	AF187066 Mus muscu
6	285.4	27.5	98274	9	HS718065	HS718065 Human DNA s
7	285.4	27.5	98274	9	HS718065	HS718065 Human DNA s
8	242.4	27.2	244	11	GI3350	GI3350 Human DNA s
9	185.2	20.8	45678	9	HS9351F8	HS9351F8 Human DNA s
10	185.2	20.8	45678	9	HS9351F8	HS9351F8 Human DNA s
11	158.6	17.8	59648	9	HS635G19	HS635G19 Homo sapi
12	141.2	15.8	1229	9	AK000959	AK000959 Homo sapi
13	141.2	15.8	1364	6	AX100231	AX100231 Sequence
14	141	15.8	43952	9	HSU105G4	HSU105G4 Human DNA s
c	132.4	14.9	477	11	G35294	G35294 human STS S
15	122.2	13.7	35714	9	HS198P4	HS198P4 Human DNA
16	117.2	13.2	154577	2	AL357046	AL357046 Homo sapi
c	117.2	13.2	154577	2	AL357046	AL357046 Homo sapi
17	117.2	13.2	154577	2	AL357046	AL357046 Homo sapi
18	117.2	13.2	154577	2	AL357046	AL357046 Homo sapi
19	117.2	13.2	154577	2	AL357046	AL357046 Homo sapi
20	110.8	12.4	41029	2	HSU80B08	HSU80B08 Human DNA
21	110.8	12.4	41029	2	HSU80B08	HSU80B08 Human DNA
22	109.8	12.3	642	9	AF183416	AF183416 Homo sapi
23	109.8	12.3	828	9	AF220189	AF220189 Homo sapi
c	109.2	12.3	40584	9	AL133348	AL133348 Human DNA
24	109.2	12.1	791	9	AF237783	AF237783 Homo sapi
25	108.2	12.1	898	6	AX078272	AX078272 Sequence
26	104	11.7	858	6	AX078272	AX078272 Sequence
c	102.4	11.5	744	9	AF251053	AF251053 Homo sapi
27	77	8.6	43952	9	HSU105G4	HSU105G4 Human DNA s
28	74.6	8.4	1163	9	BC005988	BC005988 Homo sapi
29	74.6	8.4	1163	9	BC005988	BC005988 Homo sapi
30	71.6	8.0	504	11	G24641	G24641 human STS W
31	65.4	7.5	793	10	AF097439	AF097439 Mus muscu
c	65.4	7.5	793	10	AF097439	AF097439 Mus muscu
32	59.6	6.7	40822	2	HSU017788	HSU017788 Human DNA s
33	59.6	6.7	40822	2	HSU017788	HSU017788 Human DNA s
34	59.6	6.7	138872	2	AC027560	AC027560 Homo sapi
c	59.4	6.7	176039	2	AC027560	AC027560 Homo sapi
35	59.4	6.7	176039	2	AC027560	AC027560 Homo sapi
36	58.6	6.6	1208	9	AK026349	AK026349 Homo sapi
c	58.6	6.6	91480	9	HSU823F3	HSU823F3 Human DNA
37	58.6	6.6	91480	9	HSU823F3	HSU823F3 Human DNA
c	58.4	6.6	156275	2	AL513285	AL513285 Homo sapi
38	58.4	6.6	156275	2	AL513285	AL513285 Homo sapi
39	57	6.4	196367	9	AC020740	AC020740 Homo sapi
40	56.6	6.4	2269	10	AF097437	AF097437 Mus muscu
c	56.6	6.4	168896	2	AC037486	AC037486 Homo sapi
41	56.2	6.3	127606	2	AL356238	AL356238 Homo sapi
c	56.2	6.3	127606	2	AL356238	AL356238 Homo sapi
42	56.2	6.3	127606	2	AL356238	AL356238 Homo sapi
43	56.2	6.3	127606	2	AL356238	AL356238 Homo sapi
44	56.2	6.3	127606	2	AL356238	AL356238 Homo sapi
45	56.2	6.3	127606	2	AL356238	AL356238 Homo sapi

ALIGNMENTS

RESULT	1
HUMOGC	HUMOGC 891 bp mRNA PRI 07-MAR-1995
LOCUS	Human unknown protein from clone pHR74 mRNA; complete cds.
DEFINITION	Human unknown protein from clone pHR74 mRNA; complete cds.
ACCESSION	M38188.1
VERSION	M38188.1 GI:189378
KEYWORDS	Human ovarian granulosa cell line, cDNA to mRNA.
SOURCE	Human ovarian granulosa cell line, cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 891)
JOURNAL	Rapp, G., Freudenstein, J., Klaudiny, J., Mucha, J., Hempe, F.,
MEDLINE	Zimmer, M. and Scheit, K.H.
COMMENT	Characterization of three abundant mRNAs from human ovarian granulosa cells. DNA Cell Biol. 9 (7), 479-485 (1990)
	91025550
	Draft entry and computer-readable sequence for [DNA 9, 479-485
	(1990)] kindly submitted

Qy 711 cctaataatgccttttaactgatacggtttgcgtgaaccccatatgtatttcacatgtgtcaagt 770
||| |||||||||||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db	121	CTGTATATGCTTACTGATGCATCAATTTGCTGTAACCTTATGTATGTAATTTCCATGCTGTCCAGGT	62
Qy	771	gggtctgtgttgcaggtctctctttagagattgctcttgcctcactcagtgaaagtttc	827
Db	61	GGGTGTGCTTTACCAGCTTCTTAATGGAGATGCTCTGGGACCCCAANGTTCTGTC	5
RESULT	6		
LOCUS	AF187066	700 bp mRNA	11-JUN-2000
DEFINITION	Mus musculus p75NTR-associated cell death executor	complete cds.	
ACCESSION	AF187066	GI:8452897	
VERSION	AF187066.1		
KEYWORDS	house mouse,		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;		
REFERENCE	1 (bases 1 to 700)		
AUTHORS	Mukai,J., Hachiya,T., Shoji,Hoshino,S., Kimura,M.T., Nadano,D.,		
TITLE	Suwanpo,P., Hanaka,T., Li,Y., Irie,S., Greene,L.D., and Sato,T.A.		
	NADe, a p75NTR-associated cell death executor, involved in p75NTR		
	transduction mediated by the common neurotrophin receptor		
JOURNAL	3 Biol. Chem. 275 (23), 17566-17570 (2000)		
MEDLINE	20398829		
REFERENCE	2 (bases 1 to 700)		
AUTHORS	Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suwanpo,P.,		
	Hanaka,T., Li,Y., Irie,S. and Sato,T.		
TITLE	Direct Sumoylation of p75NTR by the Ubiquitin-Protein Ligase		
JOURNAL	Direct Sumoylation of p75NTR by the Ubiquitin-Protein Ligase		
FEATURES	Location/Qualifiers		
	1..700		
	/organism="Mus musculus"		
	/db_xref="BAU/b"		
gene	/db_xref="taxon:10090"		
	1..700		
	/name="NADe"		
CDS	177..551		
	/gene="NADe"		
	/note="Involved in the common neurotrophin receptor		
	p75NTR-mediated signal transduction; NADe"		
	/codon_start=1		
	/product="p75NTR-associated cell death executor"		
	/protein_id=845288.1		
	/translation="MANVQVREHMDPVLNGSDRPPVGGEGHGPAAANNNNNNHIN		
	NHNRGGQARLLAPNFNAIPNRQNDLGGDDGDDMPMEHREIRRLRLQLRLQ		
	LRLMGSLSHHHDDHDFCLMP"		
BASE COUNT	178 a 187 c 203 g 132 t		
ORIGIN			
Query Match	35.08;	Score 312.2;	DB 10; Length 700;
Query Local Similarity	78.58;	Pred. NO. 1.7e-75;	
Matches 431; Conservative 0;	Mismatches 73;	Indels 45;	Gaps 3;
Qy	284	aaasacacccgagaaaaaatactctcatcgcgcaattattccacaggagaaagaagat	343
Db	149	AATAGACACCAAGGAAACAAATATCTCATCTGCCCATGTTCNCAGGAAACGAAGAT	208
Qy	344	ggagacacccgagaaaaaatactctcatcgcgcaattattccacaggagaaagaagat	403
Db	209	GGAGACGCCCTCGCAAGAAATGGACAGGAGACGCCCTGTGGGAGAGAGGTGAGG	268
Qy	404	gcttcgagagaa-----atcgacagggg	424
Db	269	GCTTCGTCGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	328
Qy	425	ccagctcgagagatgctgcctattctcatcgcgcaattattccacaggagaaagaagat	484

Mapping Group. Further information can be found at <http://www.sanger.ac.uk/ChGIS/>. This comparison was made by comparing clones from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EMBL, SW, SWISSPROT, Tr, TrEMBL, WPI, WormPEP, Information on the WormPEP database can be found at <http://www.wormbase.org/>. Genomic clones are identified from the human BAC library described in U-J. Kim et al. (1996) Genomics 34, 213-218.

VECTOR: pBeloBAC11

This sequence is the entire insert of clone CTA-714B7.

FEATURES

```

source
    1..98374
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /map="012-13.2"
        /clone="CTA-714B7"
        /clone_lib="CIT978SK-A2"
        /note="14 copies 2 mer ac 100 conserved"
repeat_region
    109..136
        /note="L1 repeat: matches 183..258"
repeat_region
    183..258
        /note="L1 repeat: matches 183..258"
misc_feature
    1177..1446
        /note="match: GSS: Em:AQ373724"
repeat_region
    1500..1508
        /note="L1 repeat: matches 2693..2750 of consensus"
repeat_region
    2568..5038
        /note="L1 repeat: matches 3673..6170 of consensus"
repeat_region
    5036..5486
        /note="L1 repeat: matches 2052..2207 of consensus"
repeat_region
    5496..6885
        /note="L1 repeat: matches 1..1538 of consensus"
repeat_region
    6896..7185
        /note="L1 repeat: matches 2356..2710 of consensus"
misc_feature
    7185..7611
        /note="match: GSS: Em:AQ566001"
repeat_region
    7369..7594
        /note="L1 repeat: matches 34..256 of consensus"
misc_feature
    complement(7713..8092)
        /note="match: GSS: Em:AQ071449"
repeat_region
    7904..7978
        /note="L1 repeat: matches 20..95 of consensus"
repeat_region
    8086..8300
        /note="L1 repeat: matches 8..255 of consensus"
misc_feature
    8300..8365
        /note="match: GSS: Em:AQ059599"
repeat_region
    8349..8864
        /note="L1 repeat: matches 1..280 of consensus"
misc_feature
    complement(8897..9290)
        /note="match: GSS: Em:AQ343214"
repeat_region
    9235..9361
        /note="L1 repeat: matches 2619..2749 of consensus"
misc_feature
    9378..9590
        /note="match: GSS: Em:AQ284832"
repeat_region
    10338..10358
        /note="L1 repeat: matches 1..294 of consensus"
repeat_region
    10429..10563
        /note="L1 repeat: matches 1..135 of consensus"
repeat_region
    10811..10868
        /note="L1 repeat: matches 2648..2705 of consensus"
repeat_region
    10819..11064
        /note="L1 repeat: matches 2..255 of consensus"
repeat_region
    11112..11324
        /note="L1 repeat: matches 11..255 of consensus"
repeat_region
    11868..12112
        /note="L1 repeat: matches 20..167 of consensus"

```

```

repeat_region
    12925..13223
        /note="L1 repeat: matches 1..300 of consensus"
repeat_region
    13290..13495
        /note="L1 repeat: matches 9..214 of consensus"
misc_feature
    13338..13833
        /note="match: GSS: Em:AQ568864"
repeat_region
    13688..13749
        /note="L1 repeat: matches 77..139 of consensus"
repeat_region
    13840..13966
        /note="L1 repeat: matches 22..147 of consensus"
repeat_region
    14024..14144
        /note="L1 repeat: matches 46..136 of consensus"
repeat_region
    14925..15058
        /note="L1 repeat: matches 48..191 of consensus"
repeat_region
    15994..16040
        /note="L1 repeat: matches 102..144 of consensus"
repeat_region
    16116..16184
        /note="L1 repeat: matches 3..70 of consensus"
repeat_region
    16897..17005
        /note="L1 repeat: matches 9..118 of consensus"
repeat_region
    17018..17070
        /note="L1 repeat: matches 108..192 of consensus"
repeat_region
    17748..17912
        /note="L1 repeat: matches 73..245 of consensus"
repeat_region
    18657..18871
        /note="L1 repeat: matches 35..252 of consensus"
repeat_region
    18927..19230
        /note="L1 repeat: matches 1..301 of consensus"
repeat_region
    19492..19559
        /note="L1 repeat: matches 2..67 of consensus"
repeat_region
    19560..19820
        /note="L1 repeat: matches 34..295 of consensus"
repeat_region
    19822..19959
        /note="L1 repeat: matches 67..212 of consensus"
repeat_region
    19987..20286
        /note="L1 repeat: matches 1..300 of consensus"
repeat_region
    20316..20512
        /note="L1 repeat: matches 82..262 of consensus"
repeat_region
    20513..20741
        /note="L1 repeat: matches 5970..6224 of consensus"
repeat_region
    20769..21074
        /note="L1 repeat: matches 1..303 of consensus"
repeat_region
    21085..21207
        /note="L1 repeat: matches 1776..1898 of consensus"
repeat_region
    21212..21391
        /note="L1 repeat: matches 6..178 of consensus"
repeat_region
    21392..21471
        /note="L1 repeat: matches 3..86 of consensus"
repeat_region
    22146..22441
        /note="L1 repeat: matches 1..296 of consensus"
repeat_region
    22480..22510
        /note="L1 repeat: matches 267..349 of consensus"
repeat_region
    22940..23204
        /note="L1 repeat: matches 378..650 of consensus"
repeat_region
    23205..23514
        /note="L1 repeat: matches 1..311 of consensus"
repeat_region
    23919..24194
        /note="L1 repeat: matches 36..311 of consensus"
repeat_region
    24496..24590
        /note="L1 repeat: matches 253..349 of consensus"
repeat_region
    24591..24610
        /note="L1 repeat: matches 102..143 of consensus"
repeat_region
    25378..25555
        /note="L1 repeat: matches 2330..2519 of consensus"
repeat_region
    25673..25910
        /note="L1 repeat: matches 4..255 of consensus"
repeat_region
    26722..26798
        /note="L1 repeat: matches 2572..2649 of consensus"
repeat_region
    27040..27107
        /note="L1 repeat: matches 2489..2500 of consensus"
repeat_region
    27108..27199
        /note="L1 repeat: matches 41..304 of consensus"
repeat_region
    27375..27471

```



```

RESULT 10
LOCUS   HSV351F8      45678 bp      DNA      PRI      23-NOV-1999
DEFINITION
Human DNA sequence from cosmid V351F8, between markers DXS366 and
DXS87 on chromosome X contains ESTs.
ACCESSION
270719
VERSION 270719.1 GI:1261915
KEYWORDS
X.
SOURCE  human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 45678)
AUTHORS Whiteley M.
TITLE Direct Submission
JOURNAL Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT IMPORTANT: This sequence is the entire insert of clone V351F8. The
true left end of clone V351F8 is at 1 in this sequence. The true
right end of clone V351F8 is at 45678.
FEATURES
V351F8 is from the human chromosome X-specific cosmid library.
SOURCE
1 45678
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="GHC-351F8"
/clone_lib="SCCV"
1..70
repeat_region
/partial
note="Alu repeat: matches 80..1 of consensus"
repeat_region
note="Alu repeat: matches 1..308 of consensus"
repeat_region
1406..5179
note="MSTA element fragment"
repeat_region
5243..5276
note="LI element fragment"
repeat_region
5279..5567
note="17 copies of 2 mer 82 % conserved"
/partial
note="Alu repeat: matches 304..1 of consensus"
repeat_region
9975..10369
note="MSTB element fragment"
repeat_region
10399..10701
note="MSTB element fragment"
repeat_region
11318..11578
note="THE1A8 element fragment"
repeat_region
11715..11846
note="THE1A8 element fragment"
repeat_region
12074..12170
note="MSTB element fragment"
repeat_region
12227..12486
note="MSTB element fragment"
repeat_region
12437..12502
note="MSTB element fragment"
repeat_region
12756..13059
/partial
note="Alu repeat: matches 308..1 of consensus"
repeat_region
14478..14650
note="THE1B element fragment"
repeat_region
14491..14536
note="MSTB element fragment"
repeat_region
14715..14831
note="THE1B element fragment"
repeat_region
15099..15333
note="THE1B element fragment"
repeat_region
15534..15626
note="31 copies of 3 mer 85 % conserved"
repeat_region
15566..15619
note="3 copies of 18 mer 98 % conserved"
repeat_region
16351..16436
note="LI element fragment"
repeat_region
16728..16763

```

```

repeat_region
16764..17054
/partial
note="Alu repeat: matches 308..1 of consensus"
repeat_region
19420..19449
note="15 copies of 2 mer 87 % conserved"
misc_feature
complement(19537..20013)
note="match: 3' EST NS1315 clone 283089"
repeat_region
20071..22014
note="15 copies of 2 mer 98 % conserved"
repeat_region
20073..20112
note="10 copies of 4 mer 100 % conserved"
repeat_region
20073..20126
note="3 copies of 18 mer 87 % conserved"
repeat_region
23088..23300
note="LI element fragment"
repeat_region
23311..23490
note="MSTB element fragment"
repeat_region
23318..23489
note="MSTB element fragment"
repeat_region
23502..23491
note="MSTB element fragment"
repeat_region
23507..24407
note="LI element fragment"
repeat_region
24376..24462
note="MSTC element fragment"
repeat_region
24379..24491
note="MSTB element fragment"
repeat_region
24688..24738
note="MSTB element fragment"
repeat_region
24810..25016
note="LI element fragment"
repeat_region
25284..25401
note="LI element fragment"
repeat_region
25439..25696
note="MER25 element fragment"
repeat_region
26284..26343
note="MSTB element fragment"
repeat_region
26428..26491
note="MSTB element fragment"
repeat_region
26465..26518
note="MSTB element fragment"
repeat_region
27446..27667
note="LI element fragment"
repeat_region
28567..29247
/partial
note="Alu repeat: matches 308..1 of consensus"
repeat_region
30160..30325
note="2 copies of 83 mer 98 % conserved"
misc_feature
32588..32875
note="match: 5' EST H68599 clone 239077"
misc_feature
complement(32825..33230)
note="match: 3' EST H68239 clone 289077"
repeat_region
35531..35802
/partial
note="Alu repeat: matches 1..260 of consensus"
repeat_region
35804..35839
note="18 copies of 2 mer 83 % conserved"
repeat_region
38625..38705
note="MSTB element fragment"
repeat_region
38747..38839
note="MSTB element fragment"
repeat_region
38949..39032
note="MSTB element fragment"
repeat_region
38965..39076
note="MSTB element fragment"
BASE COUNT 14260 a 3135 c 3327 g 12936 t
ORIGIN

```

```

Query Match      20.8%; Score 185.2; DB 9; Length 45678;
Best Local Similarity 64.4%; Pred. No. 5.2e-40;
Matches 345; Conservative 0; Mismatches 178; Indels 13; Gaps 4;

```



```

QY 637 tcatgcttgactctgcatcttacc 662
Db 602 TCATACCTTGAGTCTTAAAGGTTTC 627

RESULT 14
HSU105G4 43952 bp DNA PRI 23-NOV-1999
LOCUS Human DNA sequence from cosmid U105G4, between markers DXS366 and
DEFINITION DXS87 on chromosome X contains ESTs.
ACCESSION 22846
VERSION 22846.1 GI:1877217
KEYWORDS X.
SOURCE Human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 43952)
REFERENCE 1 (bases 1 to 43952)
AUTHORS Burton.J.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-1997) Sanger Centre, Hinxton, Cambridgeshire,
UK
COMMENT requests: clones requested from AC.UK
true left end of clone U105G4 is at 1 in this sequence. The true
right end of clone U105G4 is at 43952.
U105G4 is from the Lawrence Livermore National Laboratory
flow-sorted X chromosome cosmid library LLOXNC01.
FEATURES
source
1..43952
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone_lib="LLOXNC01"
/clone="LLOXNC01-105G4"
repeat_region 39..420
/notes="LIM8 repeat: matches 520..920 of consensus"
repeat_region 1121..1233
/notes="LIM1 repeat: matches 61..170 of consensus"
repeat_region 1121..1203
/notes="LIM5 repeat: matches 61..143 of consensus"
repeat_region 3153..3463
/notes="LIM5 repeat: matches 1..300 of consensus"
repeat_region 3465..3568
/notes="LIM5 repeat: matches 1..300 of consensus"
repeat_region 4207..4505
/notes="LIM5 repeat: matches 1..300 of consensus"
prim_transcript <4809..>56786
/notes="AluY repeat: matches 301..1 of consensus"
repeat_region 7275..7312
/notes="19 copies of 2 mer 84 & conserved"
repeat_region 7315..7612
/notes="AluJo repeat: matches 302..2 of consensus"
repeat_region 7661..7738
/notes="MIR repeat: matches 183..261 of consensus"
repeat_region 7881..7933
/notes="MIR repeat: matches 118..174 of consensus"
repeat_region 8074..8196
/notes="MIR repeat: matches 4965..5390 of consensus"
repeat_region 9151..10034
/notes="LIP2 repeat: matches 1..891 of consensus"
repeat_region 11622..11659
/notes="19 copies of 2 mer 100 & conserved"
repeat_region 17600..17897
/notes="AluSg repeat: matches 299..2 of consensus"
repeat_region 18288..18693
/notes="MST repeat: matches 3..424 of consensus"

```

```

repeat_region 19079..19315
/notes="MIR repeat: matches 262..13 of consensus"
repeat_region 20482..20770
/notes="AluJo repeat: matches 292..9 of consensus"
repeat_region 22479..22522
/notes="MIR repeat: matches 44..1 of consensus"
repeat_region 23553..23686
/notes="17 copies of 2 mer 97 & conserved"
repeat_region 23726..23856
/notes="AluJo repeat: matches 1..131 of consensus;
incomplete repeat"
repeat_region 23866..23958
/notes="LIM8 repeat: matches 83..1 of consensus"
repeat_region 26423..26480
/notes="29 copies of 2 mer 81 & conserved"
repeat_region 27936..28102
/notes="LIM1 repeat: matches 572..442 of consensus"
repeat_region 27978..28196
/notes="LIM2 repeat: matches 572..344 of consensus"
prim_transcript complement: 5'-TGTGTA-3'
repeat_region 29196..29370
/notes="FAM repeat: matches 175..1 of consensus"
repeat_region 29589..29640
/notes="2 copies of 26 mer 98 & conserved"
prim_transcript 30953..33085
/notes="match: multiple ESTs; match: N25344 M62822 H09953
K1228 N52068; match: N21487 H97562 H09131 T10046; match:
H09952 H01157 K14909 N31459"
misc_feature 31131..31155
/notes="Site of TN1000 transposon insertion"
repeat_region 33134..33352
/notes="LIM1 repeat: matches 304..97 of consensus"
repeat_region 34614..34742
/notes="MIR repeat: matches 101..237 of consensus"
repeat_region 35476..35566
/notes="MLT2D repeat: matches 1..88 of consensus"
repeat_region 35644..36167
/notes="MLT2D repeat: matches 75..553 of consensus"
repeat_region 36304..36907
/notes="LIM2 repeat: matches 299..1 of consensus"
repeat_region 36906..36838
/notes="LIP15 repeat: matches 870..640 of consensus"
repeat_region 36967..37400
/notes="LIM5 repeat: matches 490..913 of consensus"
repeat_region 37865..37996
/notes="LIM2 repeat: matches 285..419 of consensus"
repeat_region 37979..38231
/notes="LIM3 repeat: matches 428..698 of consensus"
repeat_region 38934..39028
/notes="LIM2 repeat: matches 472..576 of consensus"
repeat_region 39306..39463
/notes="MIR repeat: matches 235..117 of consensus"
repeat_region 39800..40160
/notes="MLT2FB repeat: matches 414..53 of consensus"
repeat_region 40760..40895
/notes="AluX repeat: matches 1..136 of consensus;
incomplete repeat"
repeat_region 40899..42656
/notes="L1 repeat: matches 3592..5390 of consensus"
repeat_region 42306..42456
/notes="LIM5 repeat: matches 1..1047 of consensus"
BASE COUNT 13852 a 8286 c 8337 g 13477 t
ORIGIN

```

```

Query Match 15.8%; Score 141; DB 9; Length 43952;
Best Local Similarity 54.0%; Pred. No. 8.6e-28;
Matches 494; Conservative 0; Mismatches 345; Indels 76; Gaps 7;
QY 14 ctcctatccgctctctcttgcctcgaagctctgggaagatcccggaac 73
Db 15173 CCCACTCCAGATCTCTCCCTTTTGTCGATAGGCTTAGGAGGATC----TACAA 15228

```

Qy 692 ttcttttctcttgatttcttctaatatgcttttactgataccggttgctgtgaaacctatg 751
Db 178 ATTTCCTGCTGCTACCTTAATGTGAGACCTTTGGCACTCTCTGCAATTTCTGATTG 119
Qy 752 ttatttccatgt 763
Db 118 GAGATTGCATTT 107

Search completed: March 11, 2002, 16:34:17
Job time: 8232 sec

USA
NIN-MGC Project URL: http://mgc.nci.nih.gov
Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: Baylor College of Medicine Human Genome
Scatter plot:
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalobebcm.tmc.edu,
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINE at: http://image.llnl.gov
Series: TAG Plate 20 Column: 6.
Location/Qualifiers
1. -2901
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:5758 IMAGE:3500522"
/tissue_type="Mammary tumor; C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI-GAP_Mam6"
/lab_name="DH105"
/cloning_vector="pCMV-SPOrt6"
755 -2359
/codon_start=1
/product="Similar to dentatorubral pallidoluysian atrophy"
/protein_id="AAH03254.1"
/db_xref="GI:13096507"
/translation="MPQQLPLTPGLPPIHVHTPLHPRAPKEHTPTFSHQSLSP
PQGLFPPSPATPQTITQIQEPASEPEPVPFPANSPKPVVDVDS
HAQSARFNHLDRGNFCNSOLFYPLEGSKLAKRADLYENREAGKADYD
YLGQPTAFETLSXARPHVASGHNNRPYVLGAIVDGILGVNVVALYSSDPAAE
REFSAREDLDRKQPKSELPKELHGVRGHRQAERVAALGNDFPLRLMLNT
YHPISLDREELAAGAFLRDMSVASLEAAERQAEVAALGNDFPLRLMLNT
PHHNEVSIHSHLLHQDDAHAASAVHYPLIDPLASGSGLTRIPVPAFTGLNFLPLP
PLHEVSRIRQLRFQAYPRDLPASLANQSAHQLOAHQAQAEQLAQLEBQQQMUIH
HPLHVSVEPDAQEDYTSHLKESKP"

CDS 623 a 958 c 791 g 529 t

alignment_scores:
Quality: 619.00 Length: 118
Ratio: 5.291 Gaps: 0
Percent Similarity: 99.153 Percent Identity: 99.153

Alignment_block:
US-09-327-750D-30 x BC003254 ..

Align seg 1/1 to: BC003254 from: 1 to: 2901

1 NetClusterAspGinglyValLyasLeuAnNetGluAnAspHi 17
223 ATGGAGTCCAAGATCAACGGCTCAAAAATCTCAACATGAGCATGCCA 272

17 sGlnLyAsGluGluGluGluGluProGlnAspThrIleArgArg 34
273 TCACAAGAGGAGGAGGAAGGAAGAACCCAGATCACCATCAGAGGG 322

34 luProlaValAlaLeuIleSerGluAlaGlyLysAsnCyAspProArg 50
323 ACCACGCTGTGGCCCCCTGACTCCGAGGTGTGGAAAAACCTGTGCCTAGA 372

51 GlyCVAsArgArpPheArgValArgGlnProIleAlaHisTrArgTr 67
|||||

/translation="MESKEKRAVNSLSEMANOENEFKSOVANKGSPALPLDAGEYC
VPRGRRERVRPILOYEDMDHMLGEPQARMRENERIGEYRQLMEKLRKQLS
HSLAVSTDPFHDDHDFCLMP"
BASE COUNT 204 a 118 c 157 g 163 t
ORIGIN

alignment_scores:
Quality: 453.50 Length: 129
Ratio: 3.943 Gaps: 3
Percent Similarity: 89.147 Percent Identity: 68.992

alignment_block:
US-09-327-750D-30 x AF183416

Align seg 1/1 to: AF183416 from: 1 to: 642

```
1 MetGluSerLysAsp...GIndlyValLysAsnLeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 ATGGAGTCCAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
16 pHisGlnLysLysGluLysGluLysProGlnAspThrIleArg 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 CAACCAAGAAATAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 93
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 rGluProAlaValAlaLeuLeuSerGluAlaGlyLysAspCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
94 GGGAGGCC...TTGGCCCTCCCTTGGATCTGCTGTAATCTGTGCT 140
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 ArgGlyIleArgArgPheArgValArgGlnProIleAlaHisTyr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
141 AGAGAAATCTAGGAGCGGTCCGCTAGCAGCCATCTCGCAGTATAG 190
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 gTropoLeuMetGlnArgValGluProGlnIleArgMetArgGlu 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
191 ATGGATATGATGATGATGATGATGATGATGATGATGATGATGATG 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 LuasValGlnArgPheGlyIleValArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 AGAATATGGAAGAGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 290
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
291 AGGAAAGCAGCTTGGATCATGCTGCGGCGCAGCAGCAGCAGCCCC 340
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
341 TCACCATGACCATCATGATGATGATGATGATGATGATGATGATGAT 377
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

seq_name: gb_pr:AF220189

seq_documentation_block:
LOCUS AF220189 828 bp mRNA PRI 04-MAY-2000
DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA,
complete cds.
ACCESSION AF220189.1 GI:7689028
VERSION AF220189.1 GI:7689028
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 828)
AUTHORS Xiao, H., Song, H., Gao, G., Ren, S., Chen, Z. and Han, Z.
TITLE A novel gene expressed in human hypothalamus
JOURNAL Unpublished
PUBLISHED
AUTHORS Xiao, H., Song, H., Gao, G., Ren, S., Chen, Z. and Han, Z.
TITLE Olfactory bulb expressed gene
JOURNAL Submitted (30-DEC-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
Location/Qualifiers

source 1..828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
CDS
167..546
/codon_start=1
/product="uncharacterized hypothalamus protein HBEX2"
/protein_id="AAF67654.1"
/db_xref="GI:7689028"
/translation="MESKEKRAVNSLSEMANOENEFKSOVANKGSPALPLDAGEYC
VPRGRRERVRPILOYEDMDHMLGEPQARMRENERIGEYRQLMEKLRKQLS
HSLAVSTDPFHDDHDFCLMP"
BASE COUNT 256 a 170 c 221 g 181 t
ORIGIN

alignment_scores:
Quality: 453.50 Length: 129
Ratio: 3.943 Gaps: 3
Percent Similarity: 89.147 Percent Identity: 68.992

alignment_block:
US-09-327-750D-30 x AF220189

Align seg 1/1 to: AF220189 from: 1 to: 828

```
1 MetGluSerLysAsp...GIndlyValLysAsnLeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 ATGGAGTCCAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 216
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
16 pHisGlnLysLysGluLysGluLysProGlnAspThrIleArg 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 CAACCAAGAAATAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 257
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 rGluProAlaValAlaLeuLeuSerGluAlaGlyLysAspCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
258 GGGAGGCC...TTGGCCCTCCCTTGGATCTGCTGTAATCTGTGCT 304
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 ArgGlyIleArgArgPheArgValArgGlnProIleAlaHisTyr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
305 AGAATATGGAAGAGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 gTropoLeuMetGlnArgValGluProGlnIleArgMetArgGlu 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 ATGGAAAGCAGCTTGGATCATGCTGCGGCGCAGCAGCAGCAGCCCC 404
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 LuasValGlnArgPheGlyIleValArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
405 AGAATATGGAAGAGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 454
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
455 AGGAAAGCAGCTTGGATCATGCTGCGGCGCAGCAGCAGCAGCCCC 504
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
505 TCACCATGACCATCATGATGATGATGATGATGATGATGATGATGAT 541
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

seq_name: gb_pr:HS198P4

seq_documentation_block:
LOCUS HS198P4 35714 bp DNA PRI 16-DEC-2000
DEFINITION Human DNA sequence from clone RP1-198P4 on chromosome Xq22 Contains
a CpG island, complete sequence.
ACCESSION AL008708
VERSION AL008708.4 GI:4826445
KEYWORDS HTG; CpG Island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 35714)
AUTHORS Corby, N.

```

alignment_scores:
  Quality: 450.50      Length: 129
  Ratio: 3.952         Gaps: 3
  Percent Similarity: 88.372  Percent Identity: 68.992

alignment_block:
  US-09-327-750D-30 x AF237783 ..

Align seg 1/1 to: AF237783 from: 1 to: 791

1 MetGluSerLysAspGln.....GlyValLysAsnLeuAsnMetGluAsnAs 16
172 ATGGAGTCCAAAGAGAACTAGCAGTAAACAGCTCTCAGCATCGGAAATGC 221
16 pHisGlnLysLysGlnGluLysGlnGluLysProGlnAspThrLeuArgA 33
222 CAACCAAGAAATCAGAGAAAGAG.....CAAGTCTCTAATTAAG 262
33 rGluProAlaValAlaLeuLysrGlnLysGluGlyLysAsnCysAlaPro 49
263 GGGAGCCC...TTGGCCCTCCCTTGGATGCTGGTGAATCTGTGTGCCT 309
50 ATGCGLYGIVAGTARGARGPheA:gValArgGlnProIleAlaHisTyrAr 66
310 AGAGAAATCTGAGGCGGTCTCCGCTGTAGCAGCCCATCTCTGCAGTATAG 359
66 gTriPAspLeuMetGlnArgValGlyGluProGlnGlnGlyArgMetArgLug 83
360 ATGAAAGATGATGCTAGCTGCTGAGAACCAACAGACGATGAGTAAAG 409
83 uGsnValGlnArgPheGlyGlyGlyAspArgGlnGlnLeuMetGluLysLeu 99
410 AGATATGCGAAGAGTATGGGAGGGGGGTGAGACAGCTGATGTAAAGCTG 459
100 ATGGuArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPR 116
460 AGGAAAGACGTTGATGTCATGCTCTGGGGGAGTCAGACATGACCCCCC 509
116 oHisHisAspHisHisAspGlnLysCysLeuMetPro 128
510 TCACCATGACCATCATGATGAGTTTGGCTTTATGCC 546

seq_name: gb_PfAF251053

seq_documentation_block:
LOCUS AF251053 744 bp mRNA PRI
DEFINITION Homo sapiens X-linked protein mRNA, complete cds.
ACCESSION AF251053
KEYWORDS MIM
SOURCE human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae
Mao Y.S., Tang R., Zhou, Z., Zhao, W., Zhao, S., Wang, M.,
Wang, X., Xie, Y., Zhou, 2., Zhou, W., Chen, X., and Wu, C.
Direct Subtraction (2004-2000). Institute of Genetics, School
of Life Sciences, Fudan University, 220 Handan Road, Shanghai
200433, China.

```



```
Em:AI929470 Em:AA319216 Em:AA317625 Em:DB1392 Em:AA317418
Em:AA366528 Em:CI5186 Em:AI291270 Em:AA160660 Em:AA612605
Em:AW160832 Em:AA384405 Em:AA160657 Em:AA091744
Em:AI291126 Em:CI5591 Em:AA384468 Em:AI291593 Em:AI320206
Em:AI929477 Em:AI326617 Em:AI706545 Em:AA464771 Em:H61106
Em:W68380 Em:AA341652 Em:AA384269 Em:AA527696 Em:W48778
Em:AA633989 Em:AA159486 Em:AW167301 Em:AI929428
Em:AA379411 Em:CI5948 Em:AI745669 Em:AW024587 Em:AW379207;
match: proteins: Tr:Q9Y516 Tr:Q9Y517*
/evidence-not_experimental
/product=jd79p11.1 (novel protein similar to mouse Bex2
(brain-expressed x-11 novel protein 2))
/gene=jd79p11.1
complement(30502..32134)
30507..30988 STS: Em:G24641; match: STS: Em:T86927*
30508..30952 match: STS: Em:R43117; match: STS: Em:G23964*
/notes=match: STS: Em:R43117; match: STS: Em:G23964*
complement(30518..30523)
/gene=jd79p11.1*
30541..30922
/notes=match: STS: Em:H61107*
31543..32262
/notes=CpG island*
32442..33170
/notes=not_experimental
complement(32655..33316)
/notes=42 copies 2 mer cc 64% conserved*
33027..34514
/notes=match: GSS: Em:AQ489303*
34528..34786
/notes=L1ME2 repeat: matches 4631..6155 of consensus*
34787..35085 repeat: matches 4404..4664 of consensus*
/notes=ALUS repeat: matches 1..300 of consensus*
35086..35412

alignment_scores:
Quality: 450.00 Length: 129
Ratio: 3.879 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667

alignment_block:
US-09-327-750D-30 x AL133348/rev ..
Align seg 1/1 to reverse of: AL133348 from: 1 to: 40584
1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||
31132 ATGGAGTCCAAAGGAGGACGCGTTAAACATCTCATCTCGAAAATGT 31083
16 phisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||
31082 CAACCGAGAAATGATGAAAGATGAAAGAGGACGATGCTGTAATAAG 31033
33 rgGluProAlaValAlaLeuIleSerGluAlaGluLysAsnCysAlaPro 49
|||||
31032 GGGAGCC...TTGGCCCTACTTGTGATGTAGTACTGTGTGCT 30986
50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||
30985 AGAGAAACCGTAGCGCTTGGCGCTTAGCGCCATCCCTCCAGTATAG 30936
66 gTPAsPLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||
30935 TTGGACATATATCATAGCTTGGAGACCCACAGCAAGATGAGAGAGG 30886
83 LuAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||
30885 AGAATATGGAAGGATGGGAGGAGGTGAGACACGTATGGGAAAGCTG 30836
100 ArgGluArgGlnLeuSerHisSerLysArgAlaValSerThrAspProPr 116
```

```
|||||
30835 AGGAAAGACAGTTGAGTCATAGTTTCGGCAGTCACGACTGATCCCC 30786
116 oHLeHLaSpHHisAspGluPheCysLeuMetPro 128
|||||
30785 TCACCATGACCATCAGCATGAGTTTTCGCTTATGCC 30749
```

seq_name: gb_hg:HSU8081

seq_documentation_block: 41029 bp DNA HTG 10-JUL-2001
LOCUS HSU8081 Homo sapiens chromosome X, clone LLOXNC01-8081 map q22.1-22.3, ***
DEFINITION Homo sapiens chromosome X, clone LLOXNC01-8081 map q22.1-22.3, ***
ACCESSION NC_000000.11
VERSION AL022169.3
KEYWORDS HTG: HTGS_PHASE2; HTGS_CANCELLED.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 41029)

REFERENCE
AUTHORS Bird, C. Substitution
TITLES Submitted (08-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 USA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

COMMENT
On Mar 12, 2001 this sequence version replaced gi:12750927.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
Project description: Project information
Center project name: CH0801

----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator ABI; 19% of reads
Chemistry: Dye-terminator Big Dye; 80% of reads
Consensus quality: 41009 bases at least Q40
Consensus quality: 41023 bases at least Q30
Insert size: 4053; sum of contigs: 41023
Insert size: 5175; 0.8% error; agarose-fp
Quality coverage: 9.67x in Q30 bases; sum-of-contigs Quality
coverage: 7.67x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1..41029 "Homo sapiens"
/dbxref="taxon:9606"
/chromosome="X"
/map="q22.1-22.3"
/clone="LLOXNC01-8081"
/clone_lib="LLOXNC01"
1..41029
misc_feature
1..41029
/note="assembly fragment:00890"

BASE COUNT 13249 a 7796 c 7781 g 12203 t
ORIGIN

alignment_scores:
Quality: 450.00 Length: 129
Ratio: 3.879 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667

alignment_block:
US-09-327-750D-30 x HSU8081 ..

Align seg 1/1 to reverse of: AL133348 from: 1 to: 40584

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||

31132 ATGGAGTCCAAAGGAGGACGCGTTAAACATCTCATCTCGAAAATGT 31083

16 phisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||

31082 CAACCGAGAAATGATGAAAGATGAAAGAGGACGATGCTGTAATAAG 31033

33 rgGluProAlaValAlaLeuIleSerGluAlaGluLysAsnCysAlaPro 49
|||||

31032 GGGAGCC...TTGGCCCTACTTGTGATGTAGTACTGTGTGCT 30986

50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||

30985 AGAGAAACCGTAGCGCTTGGCGCTTAGCGCCATCCCTCCAGTATAG 30936

66 gTPAsPLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||

30935 TTGGACATATATCATAGCTTGGAGACCCACAGCAAGATGAGAGAGG 30886

83 LuAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||

30885 AGAATATGGAAGGATGGGAGGAGGTGAGACACGTATGGGAAAGCTG 30836

100 ArgGluArgGlnLeuSerHisSerLysArgAlaValSerThrAspProPr 116

OM of: US-09-327-750D-34 to: EST.* out_format : pfe
Date: Mar 11, 2002 2:16 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compuen Ltd.

```

Command line parameters:
  -MODEL=frame.p2n.model -DEV=xlh
  -O=/cgn2/USPTO.spool/US09327750/runat_11032002_101153_20308/app_query.fasta_1.1472
  -DB=EST -QFMT=fastp -SUFFIX=rst -GAPOF=12.000 -GAPEXT=4.000
  -MINMATCH=0.100 -LOOPI=0.000 -LOOPEXT=0.000 -QGAPOF=4.500
  -CGAPEXT=0.050 -XGAPOF=10.000 -XGAPEXT=0.500 -FGAPOF=6.000
  -FGAPEXT=7.000 -YGAPOF=10.000 -YGAPEXT=0.500 -DELOF=6.000
  -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
  -ALIGN=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
  -MINLEN=0 -MAXLEN=400000000 -USER=US09327750_ACGN_1_3554
  -NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
  -NO_XLUPX1 -WAIT_THREADS=1

Search information block:
  Query: US-09-327-7500-34
  Query length: 118
  Database: EST.*
  Database sequences: 11351937
  Database length: 1077921985
  Search time (sec): 4085.940000

```

sequence list	Strd	Orig	zScore	EScore	Len	Documentation
gb.est1:A1278667	-	520.00	1031.44	2.76-48	568	A1278667 EST224562 Normalized
gb.est1:A1289546	-	592.00	984.02	1.26-45	643	B1289546 UT1-ER-DKO-off-c-12-0-UT
gb.est1:A1289546	-	546.00	906.12	2.66-41	789	A1289546 UT1-mus musculus ES cells
gb.est1:A1304040	-	546.00	906.10	2.66-41	792	A1304040 Mus musculus ES NIA Mouse
gb.est1:BG0808261	-	534.00	889.95	2.16-40	458	BG0808261 UT1-Harvard Mammalian
gb.est1:A137325	-	534.00	889.95	2.16-40	458	A137325 UT1-Harvard Mammalian
gb.est1:AM231884	-	520.00	865.41	4.86-35	717	AM231884 UT1-Harvard Mammalian
gb.est1:BG0790271	-	503.00	834.94	1.86-37	519	BG0790271 UT1-Harvard Mammalian
gb.est1:BG133771	-	483.00	804.95	1.18-35	463	BG133771 us27h062 UT1 Soares NMH
gb.est1:W17172	-	466.00	779.34	3.18-34	317	W17172 m77e12 UT1 Soares NMH
gb.est1:A1117429	-	444.00	742.75	3.36-32	312	A1117429 m77e12 UT1 Soares NMH
gb.est1:A1012535	-	436.00	710.42	2.16-30	444	A1012535 EST5206986 Normalized
gb.est1:A1058693	-	435.00	708.24	2.76-30	472	A1058693 mC5390 UT1 Soares mouse
gb.est1:W54487	-	410.00	683.49	6.66-39	465	W54487 mD05931 UT1 Soares mouse
gb.est1:A1118340	-	407.00	679.16	1.26-38	422	A1118340 v9112026 Beddington mouse
gb.est1:BF722027	-	406.00	680.44	1.06-38	286	BF722027 m27e12 UT1 Soares NMH
gb.est1:A1278667	-	398.00	658.52	6.58-37	442	A1278667 m27e12 UT1 Soares NMH
gb.est1:A1278667	-	386.00	644.71	9.56-37	393	A1278667 m5902 UT1 Soares NMH
gb.est1:AAG93925	-	395.00	641.68	1.48-36	467	AAG93925 EST197728 Normalized
gb.est1:BF556502	-	373.00	623.91	1.48-25	348	BF556502 UT-ER-f1-f07-0-UT
gb.est1:A1517301	-	372.00	618.13	2.96-25	639	A1517301 UT1-ER-f1-f07-0-UT
gb.est1:BE884414	-	372.00	617.30	3.26-25	722	BE884414 UT1-ER-f1-f07-0-UT
gb.est1:A1541424	-	372.00	616.23	3.76-25	845	A1541424 UT1-ER-f1-f07-0-UT
gb.est1:A1117899	-	372.00	616.08	3.86-25	865	A1117899 UT1-ER-f1-f07-0-UT
gb.est1:BG145120	-	372.00	615.77	3.96-25	905	BG145120 UT1-ER-f1-f07-0-UT
gb.est1:BG079027	-	372.00	615.59	4.06-25	929	BG079027 602825748F1 NIH MGC
gb.est1:BG035675	-	372.00	615.36	4.26-25	975	BG035675 602825745F1 NIH MGC
gb.est1:A1530650	-	372.00	615.21	4.26-25	983	A1530650 UT1-ER-f1-f07-0-UT
gb.est1:A1533004	-	372.00	614.94	4.36-25	1023	A1533004 UT1-ER-f1-f07-0-UT
gb.est1:A15349875	-	372.00	614.80	4.46-25	1044	A15349875 UT1-ER-f1-f07-0-UT
gb.est1:BF930215	-	367.00	611.31	6.96-25	579	BF930215 IL5-N70027-111200-319
gb.est1:A1534630	-	367.00	610.64	7.56-25	566	A1534630 UT1-ER-f1-f07-0-UT
gb.est1:A1532630	-	365.00	603.28	1.86-24	543	A1532630 KCO-1700001-761199-011
gb.est1:BG327431	-	364.00	603.54	1.96-24	772	BG327431 602557227F1 NIH MGC
gb.est1:BG079027	-	362.00	602.97	2.06-24	581	BG079027 UT1-GN0319-021200-326
gb.est1:BG090563	-	362.00	597.32	4.26-24	1182	BG090563 UT1-GN0319-021200-326
gb.est1:BG104887	-	360.00	600.82	2.76-24	489	BG104887 QV3-N70078-1120201-502

COMMENT

```

FEATURES
  source
    Location/Qualifiers
      1..458
      /db_xref="taxon:10090"
      /db_xref="IMAGE:872056"
      /clone_lib="Barstead MPLRB1"
      /sex="mixed"
      /tissue_type="Kidney"
      /dev_stage="6 weeks"
      /lab_host="DH10B"
      /note="Vector: p773D-Pac (Pharmacia) with a modified oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the Not I and Eco RI sites of the modified p773 vector. Library constructed by Bob Barstead."
      147 a 88 c 115 g 108 t
BASE COUNT
ORIGIN
  alignment_scores:
    Quality: 534.00 Length: 116
    Ratio: 4.944 Gaps: 0
    Percent Similarity: 93.103 Percent Identity: 86.207
  alignment_block:
    US-09-327-750D-34 x AA473525
  Align seg 1/1 to: AA473525 from: 1 to: 458
  3 SerLysValLysGlnValIleLeuAspLeuThrValGluLysAspLysLy 19
  13 GCCAAATTTAAACAAGTCATATCTGCTGATCTGCTGGAGAACAAACAAA 62
  19 sAnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGluSerHisH 36
  63 AGACAAAAGGTGGAGAGCGCTCCANCAAGTGAAGAGAACCCCAACC 112
  36 IsLeuGluGluValGluLysenLysProGlyGlyAsnValArgLys 52
  113 ATCTGGAGAGCTTGAACAAAGAGCTGGGGGAATGTCGGAAGGAAA 162
  53 ValArgArgLeuValProAsnPhelLeuTrpAlaIleProAsnArgHisVa 69
  163 CTCAGGCGACTTGTGCTTAACCTTCTCTGGCCATACCAAAATAGCATGT 212
  69 IsAspHisSerGluGlyGlyGluValGlyArgPheValGlyGlnValM 86
  213 TATTCGCAATGAGGGAGAGATGTTGGAGATTGTATGTCAGGGAA 262
  86 cTGluaLysArgHisSerLysGluGlnGlnMetArgProTyrThrArg 102
  263 CAGAGTTCAGAGAAAGACTACGAGCAGCAGGAGGCTTACAGCGT 312
  103 PheArgThrProGluProAspAsnHisTyrAspPheCysLeuIlePro 118
  313 TTCGAGACCCCGAAGCTGCACATCATCTAGCATTGCTCATACCT 360
  seq_name: gb_est1:AW251884
  seq_documentation_block:
    LOCUS AW251884 540 bp mRNA 17-DEC-1999
    DEFINITION UI-R-BJ0-adn-b-06-0-UI.g1 UI-R-BJ0 Rattus norvegicus cDNA clone
    ACCESSION AW251884
    VERSION AW251884.1 GI:6595475
    KEYWORDS EST, cDNA
    ORGANISM Rattus norvegicus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
  AUTHORS
  TITLE
  JOURNAL
  MEDLINE
  COMMENT
  1 (Bases 1 to 540)
  Rattus
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Bonaldo,M.F., Lennon,G. and Soares M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
  97044477
  Contact: Soares, MB
  Program for Rat Gene Discovery and Mapping
  451 Eckstein Medical Research Building Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: msoares@blue.weeg.uiowa.edu
  The sequence contained an oligo-dT track that was present in the
  oligonucleotide that was used to prime the synthesis of first
  strand cDNA and therefore this may represent a bonafide poly A
  tail. The sequence tag present in the cDNA between the Not I site
  and the oligo-dT track is 16.5 dpc library cDNA library Preparation:
  M.B. Soares Lab Clone distribution: clones will be available
  through Research Genetics (www.resgen.com)
  Seq primer: M13 Forward
  POLYA=Yes.
  Location/Qualifiers
    1..540
    /organism="Rattus norvegicus"
    /strain="Sprague-Dawley"
    /db_xref="taxon:10090"
    /clone_lib="UI-R-BJ0-adn-b-06-0-UI"
    /clone_lib="UI-R-BJ0"
    /dev_stage="adult"
    /lab_host="DH10B (Life Technologies)"
    /note="Vector: p773D-Pac (Pharmacia) with a modified oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the Not I site and the oligo-dT track is 16.5 dpc library cDNA library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
    6: 791-806, 1996.
    TAG_L1B=UI-R-BJ0
    TAG_TISSUE=atrium at 16.5 dpc
    TAG_VECTOR=CHTC
    94 g 180 t
  alignment_scores:
    Quality: 520.00 Length: 97
    Ratio: 5.417 Gaps: 0
    Percent Similarity: 98.969 Percent Identity: 98.969
  alignment_block:
    US-09-327-750D-34 x AW251884/rev
  Align seg 1/1 to reverse of: AW251884 from: 1 to: 540
  22 LysGlyGlyLysAlaSerLysGlnSerGluGluGluSerHisHisLeuG1 38
  538 AAAGTGGGAGGCTCCAAACAAAGTGAAGAGGATCCCAACCTCTGGA 489
  38 uGluValGluAsnLysProGlyGlyAsnValArgLysValArgLys 55
  488 AGAGGTTCAGAACACAGAGCTGGGGCAATCTCAGGAGGAAGTCAGGC 439
  55 rGluValProAsnPhelLeuTrpAlaIleProAsnArgHisValAspHis 71

```


/clone="IMAGE:538606"
 /clone_lib="beddingston mouse embryonic region"
 /sex="pooled"
 /tissue_type="embryo"
 /dev_stage="7.5dpc"
 /lab_host="DH12S"
 /note="Organ: whole embryo; Vector: pCMV-SPORT; Site: 1: Salt; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Gastrulating embryos were collected at 7.5dpc from C57BL/6 x DBA matings, excluding embryos that had developed solid yolk and extraembryonic tissues. 1 head stage and 3 bb (range 0.5 - 3.0 kb). Referenced in Development 121, 2479-2489 (1995)".
 BASE COUNT 110 a 56 c 92 g 54 t
 ORIGIN

alignment_scores:
 Quality: 444.00 Length: 100
 Ratio: 4.774 Gaps: 0
 Percent Similarity: 93.000 Percent Identity: 86.000

alignment_block:

US-09-327-750D-34 x AAL17429 ..

Align seg 1/1 to: AAL17429 from: 1 to: 312

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
 |||||
 12 ATGCATCCCAATTAACAAAGTCATCTGATCTCATGTGAGAAAGA 61
 |||||
 17 pLysValProAsnPheLeuTrpAlaIleProAsnArgHisValAspHis 34
 |||||
 62 CAAAGAGCAAAAAGGTGGAGGCGCTCAACAAAGTGAAGAGAAC 111
 |||||
 34 erHisLeuGluGluValGluAsnValSerProGluValAsnValArg 50
 |||||
 112 CCACCATCTGGAGAGGTGGAACAAGAGCTGGGGAAATCTCCGA 161
 |||||
 51 ArgValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
 |||||
 162 AGCAAGTCAGCGCACTTGTGCTTAACCTTCTCTGGGCATACCTAATAG 211
 |||||
 67 gHisValAspHisSerGluGlyGluGluValGluValArgPheValGly 84
 |||||
 212 GCATGTTCATCCCAATGAAGGAGGAGAGGTGTGGAGATTTGTAGTGC 261
 |||||
 84 lnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyr 100
 |||||
 262 AGGCAACAGAGTCAGAAAGACTACGAGCAGCAGCGTGGGCGCTTAC 311
 |||||

seq_name: gb_est1:A1012535

seq_documentation_block:
 LOCUS A1012535 444 bp mRNA EST 15-JUN-1998
 DEFINITION EST206986 Normalized rat placenta, Bento Soares Rattus sp. cDNA
 clone RPLAY54 3' end, mRNA sequence.

ACCESSION A1012535.1 GI:3226367

VERSION A1012535

KEYWORDS Rattus sp.

SOURCE Rattus sp.

ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 444)

AUTHORS Lee N.H., Glodok A., Chandra I., Mason T.M., Quackenbush J.,

Kerlavage A.R. and Adams M.D.

TITLE Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat

Gene Index

JOURNAL Unpublished (1998)

COMMENT Submitted to NCBI

CONTACT The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-3529
 Fax: (301)-838-0208
 Email: nhlee@tigr.org
 Seq primer: M13-21.

FEATURES

source
 1..444
 /organism="Rattus sp."
 /db_xref="taxon:10118"
 /clone_lib="Normalized rat placenta, Bento Soares"
 /note="Organ: placenta; Vector: pTT73pac; Site: 1: EcoRI;
 Site 2: NotI.
 BASE COUNT 135 a 104 c 81 g 124 t
 ORIGIN

alignment_scores:
 Quality: 426.00 Length: 80
 Ratio: 5.462 Gaps: 0
 Percent Similarity: 97.500 Percent Identity: 97.500

alignment_block:

US-09-327-750D-34 x A1012535/rev ..

Align seg 1/1 to reverse of: A1012535 from: 1 to: 444

39 GluValGluAsnLysPheGlyGlyAsnValArgLysValArgAr 55
 |||||
 444 GAGGTGAAACATGAAAGCTGGGGCAATGTCAGAGGAAAGTCAGGCG 395
 |||||
 55 gLeuValProAsnPheLeuTrpAlaIleProAsnArgHisValAspHis 72
 |||||
 394 ACTGTGCTTACTTCTATAGGCCATCTATAGGCATGTGATCAACA 345
 |||||
 72 erGluGlyGluGluValGlyArgPheValGlyGluValMetGluAla 88
 |||||
 344 GTGAAGGGGAGAGAGGTGGGAGATTCGTAGGCGAGGTGATGGAAGCC 295
 |||||
 89 LysArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgH 105
 |||||
 294 AAGAAAGTCTTAAGAGCAACAGATGAGGCGCTTACACGCGTTCCGAA 245
 |||||
 105 rProGluProAsnHisSerLysPheCysLeuIlePro 118
 |||||
 244 CCTGACCTGACCAATCATACGACTTTGGCTCTACCT 205
 |||||

seq_name: gb_est2:W35893

seq_documentation_block:
 LOCUS W35893 472 bp mRNA EST 14-MAY-1996
 DEFINITION mc53909.r1 Soares mouse embryo NM013.5 14.5 Mus musculus CDNA
 clone IMAGE:352288 5' similar to SW:HG74_HUMAN 000994 OVARIAN
 GRANULOSA CELL 13.0 KD PROTEIN HG74. [1] ;, mRNA sequence.

ACCESSION W35893

VERSION W35893.1 GI:1317733

KEYWORDS EST.

SOURCE House mouse.

ORGANISM Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 472)

AUTHORS Marzai, Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Rat Genome Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

human

Tr

11(mouse)

```

polyA_signal      790..795
/note="putative"
polyA_site        809
/note="putative"
BASE COUNT       208 a 211 c 227 g 163 t
ORIGIN
alignment_scores:
  Quality: 688.00      Length: 124
  Ratio: 5.548         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 98.387
alignment_block:
US-09-327-750D-12 x AK004531 ..
Align seg 1/1 to: AK004531 from: 1 to: 809
1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
181 ATGGCCAAATGTCACCAAGAAACAGAGAGATGGAGCCCTTCGAGAA 230
17 nGlyGluGluAspArgProValGlyGlyGlyGlnProAlaG 34
231 TGGACAGAGAGACCCCTGTGGGAGAGGTGAGGCGCCACGAGCTGCTG 280
34 lyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArgArg 50
281 CAAACACACACACACACACACACACACACACACACACACACACAC 330
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTTPalaIleProAsnAr 67
331 GCCCAGGCTGCCGAGCTTGCCTTAACCTCCGATGGGCAATCCCAACAG 380
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
381 GCAGATGAATACAGCGGTGGTGGGAGATGGAGATGATGGAAATGTCA 430
84 etGluGluMetArgGluIleArgArgGlyLeuArgGluLeuGlnLeuArg 100
431 TGGAGGAGATGACAGAGATCCCGAGAAAGCTTAGGAGCTACACCTGAGA 480
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAsnHisAspHisH 117
481 AATTGTCTACCATCTTATGGGGAGCTGTCTAACCAACACGATCACCAC 530
117 sAspGluPheCysLeuMetPro 124
531 TGTGAAATTCGCTTATGCTT 552
seq_name: gb_est1:AA253897
seq_documentation_block:
LOCUS      AA253897      468 bp      mRNA      EST
DEFINITION my31e10.r1 Barstead mouse pooled organs MFLRB4 Mus musculus cDNA
            clone IMAGE:697482 5' similar to gb:M38188 OVARIAN GRANULOSA CELL
            13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION  AA253897
VERSION    AA253897
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus.
REFERENCE  1 (bases 1 to 468)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubouque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,K., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE      The WashU-HHMI Mouse EST Project
JOURNAL    Unpublished (1996)
COMMENT    Contact: Marra M/Mouse EST Project

```

WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8301, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1800
Email: mouseest@wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:431042
Seq primer: -28m13 rev2 RT from Amersham
High quality sequence stop: 455.
Location/Qualifiers
1..468

FEATURES

```

source
1..468
/organism="Mus musculus"
/strain="FVB/N" 10090"
/db_xref="taxon:10090"
/clone="IMAGE:697482"
/seq="m13" Barstead mouse pooled organs MFLRB4"
/seq="m13" Barstead mouse pooled organs MFLRB4"
/tissue_type="pooled organs"
/dev_stage="7 day"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a NotI - oligo(dT) primer [5'
TGTACGAATCTGAAGTGGAGCGGCGCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(GGATTCGATC) and digested with NotI and ClaI to create
the cDNA library. The cDNA was then ligated into the modified
pT73 vector.
Library constructed by Bob Barstead."
BASE COUNT      140 a 126 c 120 g 82 t
ORIGIN

```

```

alignment_scores:
  Quality: 687.00      Length: 124
  Ratio: 5.540         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 97.581
alignment_block:
US-09-327-750D-12 x AA253897 ..
Align seg 1/1 to: AA253897 from: 1 to: 468

```

```

1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
49 ATGGCCAAATGTCACCAAGAAACAGAGATGGAGCCCTTCGAGAA 98
17 nGlyGluGluAspArgProValGlyGlyGlyGlnProAlaG 34
99 TGGACAGAGAGACCCCTGTGGGAGAGGTGAGGCGCCACGAGCTGCTG 148
34 lyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArgArg 50
149 CAAACACACACACACACACACACACACACACACACACACACAC 198
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTTPalaIleProAsnAr 67
199 GCCCAGGCTGCCGAGCTTGCCTTAACCTCCGATGGGCAATCCCAACAG 248
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
249 GCAGATGAATACAGCGGTGGTGGGAGATGGAGATGATGGAAATGTCA 298
84 etGluGluMetArgGluIleArgArgGlyLeuArgGluLeuGlnLeuArg 100
299 TGGAGGAGATGACAGAGATCCCGAGAAAGCTTAGGAGCTACACCTGAGA 348
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAsnHisAspHisH 117
349 AATTGTCTACCATCTTATGGGGAGCTGTCTAACCAACACGATCACCAC 398
117 sAspGluPheCysLeuMetPro 124

```

102(b)
11(mouse)

```
1. 612
source
/organism="Mus musculus"
```

11(mouse)

```

US-09-327-750d-12 x W46041
Align seg 1/1 to: W46041 from: 1 to: 492

1 MetAAsnValHisGlnGluAsnGluMetGluInProLeuGlnAs 17
|||||
13 ATGGCCATGTCCTCCACGAAAGACGAGATGGAGCCCTCGAGAA 62
|||||
17 nGlyGluGluAspArgProValGlyGlyGlyHisGlnProAla 34
|||||
63 TGGAGAGGAGAGAGCCCTGTGGGAGGAGGTGAGGCCACCCCTG 112
|||||
34 lYAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHis 112
|||||
113 CAAACACACACACACACACACACACACACACACACACACAC 162
|||||
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTAlaIlePro 67
163 GCGCAGGCTGCGCAGCTTCCCTTACTTCCGATGGCCATCCCA 212
|||||
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMet 84
|||||
213 GCAGATGAATGACGGGTGGGTGGAGATGGAGATATGGAATGT 262
|||||
84 etGluGluMetArgGluLeuArgGlyLeuArgGluLeuGlnLe 100
|||||
263 TGGAGAGATGAGAGAGATGCGGAGAGAGCTTAGGAGCTACAG 312
|||||
101 AsnCysLeuArgGlyLeuMetGlyLeuSerAsnHisAsnHis 117
|||||
313 AATTGCTACGATCTTATGGGGAGCTCTTACACACACACAC 362
|||||
117 sAspGluPheCysLeuMetPro 124
363 TGATGAATTCGCTTATGCTT 384

seq_name: gb_est1.A1006575
seq documentation block: 538 bp mRNA EST 12-JUN-1998
LOCUS A1006575 Mus musculus Mus musculus
DEFINITION uel506.y1 Sugano mouse embryo mews Mus musculus
IMAGE:1480450 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN); mRNA sequence.
ACCESSION A1006575
VERSION A1006575.1 GI:3126184
KEYWORDS EST.
SOURCE House mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 538); Murakami, Y., et al. 1998.
AUTHORS Marita, M., Hillier, L., Allen, M., Boules, M., Dietrich, N., Dubuque, T.,
Gaisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMNI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marka M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LIND; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:928806
Seq primer: primer name ambiguous
High quality sequence stop: 433.
Accession: A1006575
Source: Locus/Qualifiers
1: 538
/organism="Mus musculus"

```

```

/strain="C57BL"
/submitter="taxon:100900"
/submitter="taxon:100900"
/clone_lib="Sugano mouse embryo mews"
/dev stage="embryo, 14 dpc"
/lab host="Dh10B"
/notes="Site 1: DrIII (CACTGTGTG); Site 2: DrIII (CACTGTGTG);
with an oligo(dT) primer (ATGTGCTTTT)
double-stranded cDNA was ligated to a DrIII adaptor
[GTGGGCTACTGG], digested and cloned into distinct DrIII
sites of the pMT8S-F13 vector (5' site CACTGTGTG, 3' site
CACTGTGTG). Size of the cDNA fragments was determined by
Agarose gel electrophoresis. Library constructed by
(University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTAAAGCTGCG and 3' end primer
CGACCTGCGAGCTCGACACA."
BASE COUNT 154 a 140 c 162 g 82 t
ORIGIN

alignment_scores:
Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:
US-09-327-750d-12 x A1006575
Align seg 1/1 to: A1006575 from: 1 to: 538
1 MetAAsnValHisGlnGluAsnGluMetGluInProLeuGlnAs 17
|||||
162 ATGGCCATGTCCTCCACGAAAGACGAGATGGAGCCCTCGAGAA 211
|||||
17 nGlyGluGluAspArgProValGlyGlyGlyHisGlnProAla 34
|||||
212 TGGAGAGGAGAGAGCCCTGTGGGAGGAGGTGAGGCCACCCCTG 261
|||||
34 lYAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHis 50
|||||
262 CAAACACACACACACACACACACACACACACACACACACAC 311
|||||
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTAlaIlePro 67
|||||
312 GCGCAGGCTGCGCAGCTTCCCTTACTTCCGATGGCCATCCCA 361
|||||
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMet 84
|||||
362 GCAGATGAATGACGGGTGGGTGGAGATGGAGATATGGAATGT 411
|||||
84 etGluGluMetArgGluLeuArgGlyLeuArgGluLeuGlnLe 100
|||||
412 TGGAGAGATGAGAGATCCGAGAGAGCTTAGGAGCTACAGTGA 461
|||||
101 AsnCysLeuArgGlyLeuMetGlyLeuSerAsnHisAsnHis 117
|||||
462 AATTGCTACGATCTTATGGGGAGCTGTCTAACACACACAC 511
|||||
117 sAspGluPheCysLeuMetPro 124
512 TGATGAATTCGCTTATGCTT 533

seq_name: gb_est1.BE334866
seq documentation block:
LOCUS BE334866 542 bp mRNA EST 14-JUL-2000
DEFINITION u90a11.y1 Soares mammary gland NLMG Mus musculus cDNA clone
IMAGE:3325628 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN); mRNA sequence.
ACCESSION BE334866

```

11 (mouse)

101 AnCysLeuArgIleLeuMetGlyGluLeuSerAnshHshAspHisH 117
|||||.....
478 AATGTCTACCGATCTTATGGGGAGCTGTCTAACCAACGATGACCA 527
|||||.....
117 sAspGluPheCysLeuMetPro 124
528 TGATGAATTCGCTTATGCTT 549
|||||.....

seq_name: gb_estl.A1152323

seq_documentation_block: 430 bp mRNA 30-SEP-1998
LOCUS A1152323
DEFINITION ud79a02.r1 Soares mammary_gland_NMLMG Mus musculus
IMAGE:1477034.5, similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN MGR74 (HUMAN);, mRNA sequence.

A1152323.1 GI:3680792

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

MUS MUSCULUS

house mouse

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Mus musculus

Marx W., Hill D., Allen M., Boyles M., Dietrich N., Dubuque T.,

Geisel S., Kucaba T., Lacy M., Le M., Martin J., Moore M.,

Schellenberg K., Steptoe M., Tan F., Underwood K., Woode B.,

Theisinger B., Wylie T., Lennon G., Soares B., Wilson R. and

Waterston R.

The Nashville Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

Washington University School of Medicine

744 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1810

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:925390

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 419.

Location/Qualifiers

1..430

/organism="Mus musculus"

/clone="IMAGE:1477034"

/sex="female (lactating)"

/tissue_type="mammary gland"

/lab_host="DH10B"

/note="vector: pT73D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from mammary

gland tissue from a lactating female, and was then primed

with a Not I - oligo(dT) primer. Double-stranded cDNA was

ligated to Eco RI adaptors (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the

mod vector. Library is normalized. Library

was constructed by Benito Soares and M. Fatima Bonaldo."

135 a 111 c 113 g 71 c

BASE COUNT

ORIGIN

source

alignment_scores:

Quality: 681.00 Length: 124

Ratio: 5.537 Gaps: 0

Percent Similarity: 99.194 Percent Identity: 97.581

alignment_block:

US-09-327-750D-12 x A1152323

Align seq 1/1 to: A1152323 from: 1 to: 430

Tue Mar 12 09:01:15 2002

Clms 1, 2, 3, 4, 11 (mouse)¹³, 20, 24, 22, 23

```

34 lyranetaspnanbenasnoHishanhihsnshisHishisRatG 50
|||||
277 CAAACAACACAAACACCAACCACCAACCAACCAACCAACGAAGA 326
|||||
51 GtYGINAlAaTgArGLeUaLaPfoAsnPheArgTrpAlaIlePtoAsnAr 67
|||||
327 GCCCAGGCCTGCCGACTTGCCCTTAATCTCGATGGCCCATTCCCAACAG 376
|||||
67 ggInMetaAsnApSpLyauedgYcLYaSPGiYasPhapSpatGMetPhom 84
|||||
377 GCAGATAGATACGGGTGTTGGTGGAGATGGAGATGATGATGAATATTCCA 426
|||||
84 exGLUGuetaTargLULlAaTgArGLyLeuArgLyuGLuGLinLeuArg 100
|||||
427 TGSAGGAGATGAGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTGAGA 476
|||||
101 AsnCysLeuAaTgiLeuMeTgLYGLuLeuSerAsnHisHiasPhISH 117
|||||
477 AATTTCTAGCATCTTATGGGAGCTGCTTACCACCAACAGTACCA 526
|||||
117 asPLglUpheCYsteuWepTo 124
|||||
527 TGATGAATCTGCCTTATGCCT 548
|||||

seq_name: gb_to:AF097440

seq_documentation_block: 854 bp mRNA ROD
AF097440 Mus musculus brain expressed X-linked protein 3 [transcript]
Complete cds.
AF097440.1 GI:4580593 house mouse.
MUS MUSCULUS house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.; Mus. 1 (bases 1 to 854) F. Brown, A.L. and Kay, G.F. is a member of a novel gene family on the mouse X chromosome Hum. Mol. Genet. 8 (4), 611-619 (1999) 99172070 Erratum:[published erratum appears in Hum Mol Genet 1999 May;8(5):943]] 2 (bases 1 to 854) Brown, A.L. and Kay, G.F. Direct Submission (1998). Cancer Unit, Queensland Institute of Medical Research, Herston Rd, Brisbane, Qld 4029, Australia Location/Qualifiers
1..854 /organism="Mus musculus"/strain="FVB/N"/db_xref="taxon:10090"/db_xref="dbEST:AAZ72375"/chromosome=X/map=near Flp/label=FLP-type-pooled organs"171854
1 /gene="Bex3"172..546
1 /codon_start=1/product="brain expressed X-linked protein 3"/protein_id="AA024431.1"/db_xref="GI:4580594"/translation="MTNFWAVKFNKNKGDLGGDGDMMPEHWIRKLRLQLDRNCATLRIILGHSHHDHFCECLMP."
177 t

BASE COUNT
237 a 212 c 228 g 177 t

```

7

11(moluc)

```

US-09-327-750D-12 x BC088461
Align seg 1/1 to: BC088461 from: 1 to: 599

1 MetAlaAsnValHisGlnGlnAsnGluMetGluGlnProLeuGlnAs 17
175 ATGCCATATGTCACACAGAAAGAGAGATGGACACCCCTGCGAGA 224
17 nGlyGluGluAspArgProValGlyGlyGlyHisGlnProAlaG 34
225 TGGACAGGAAGACCCCTCTGGAGGAGGTGAGGCGCCACAGCTGCTG 274
34 lYanAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArgArg 50
275 CAACACACACACACACACACACACACACACACACACACACACAG 324
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAlaLeuProAsnAr 67
325 GCCCAGGCTGCCGACCTTGCCTTACCTTCGATGGCCCTTCCACAG 374
67 gGlnMetAsnAspGlyLeuGlyGlyGlyAspGlyAspMetGluMetPheM 84
375 GCAGATGAATGACGGTGGTGGAGATGGAGATGATGGAATGTTC 424
84 etGluGluMetArgGluLeuArgGlyLysLeuArgGluLeuGlnLeuArg 100
425 TGGAGAGATGGAGAGATGGAGAGAGCTTGGAGGAGCTACAGCTGAGA 474
101 AsnCysLeuArgTleLeuMetGlyGluLeuSerAsnHisAsnHisAsnHis 117
475 AATTGTCTACCATCTTATGGGGAGCTGTCTTAACCAACCAAGATCACA 524
117 sAspGluPheCysLeuMetPro 124
525 TGATGATCTGCTTATGCT 546

seq_name: gb_est1-AA272375
seq_documentation_block:
LOCUS AA272375 612 bp mRNA EST 26-MAR-1997
DEFINITION vb62907.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
clone IMAGE:761628 5' similar to gb:M38188 OVARIAN GRANULOSA CELL
13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION AA272375
VERSION AA272375.1 GI:1910706
KEYWORDS
SOURCE house mouse
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
1 (bases 1 to 612)
Marré,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Weisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
The WashU-HMIT Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMIT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
NOTE: This clone is available royalty-free through LNL; contact the
IMAGE Consortium (linfo@image.lnl.gov) for further information.
NCI:462548
Seq primer: -28ml3 rev2 ET from Anersham
Seq quality sequence stop: 507.
Location/Qualifiers
1..612
/organism="Mus musculus"

```

Tue 4

```

/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="Barstead mouse pooled organs MPLRB4"
/sex="mixed"
/tissue_type="pooled organs"
/dev_stage="7 day"
/lab_host="BHI08"
/notes="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a double-stranded polylinker; Site: 1: EcoRI; Site: 2: NotI; 1st
strand cDNA was ligated to EcoRI and NotI adaptors; 5'
TGTACGATCTGAGTGGAGCGCGCCCTTGTATTTTGTATTTTGTATTTT
3': double-stranded cDNA was ligated to EcoRI adaptors
(GTGGATTCGCTACC), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."
BASE COUNT 168 a 167 c 175 g 102 t
ORIGIN

alignment_scores:
Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:
US-09-327-750D-12 x AA272375
Align seg 1/1 to: AA272375 from: 1 to: 612

1 MetAlaAsnValHisGlnGlnAsnGluMetGluGlnProLeuGlnAs 17
175 ATGCCATATGTCACACAGAAAGAGATGGACACCCCTGCGAGA 225
17 nGlyGluGluAspArgProValGlyGlyGlyHisGlnProAlaG 34
226 TGGACAGGAAGACCCCTCTGGAGGAGGTGAGGCGCCACAGCTGCTG 275
34 lYanAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArgArg 50
276 CAACACACACACACACACACACACACACACACACACACACAG 325
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAlaLeuProAsnAr 67
326 GCCCAGGCTGCCGACCTTGCCTTACCTTCGATGGCGCATTCACCAACAG 375
67 gGlnMetAsnAspGlyLeuGlyGlyGlyAspGlyAspMetGluMetPheM 84
376 GCAGATGAATGACGGTGGTGGAGATGGAGATGATGGAATGTTC 425
84 etGluGluMetArgGluLeuArgGlyLysLeuArgGluLeuGlnLeuArg 100
426 TGGAGAGATGGAGAGATGGAGAGAGCTTGGAGGAGCTACAGCTGAGA 475
101 AsnCysLeuArgTleLeuMetGlyGluLeuSerAsnHisAsnHisAsnHis 117
476 AATTGTCTACCATCTTATGGGGAGCTGTCTTAACCAACCAAGATCACA 525
117 sAspGluPheCysLeuMetPro 124
526 TGATGATCTGCTTATGCT 547

seq_name: gb_est2:W81757

```

```

seq_documentation_block:
LOCUS W81757 616 bp mRNA EST 12-SEP-1996
DEFINITION me95d06.r1 Soares mouse embryo NMEL13.5 14.5 Mus musculus cDNA
clone IMAGE:403307 5' similar to gb:M38188 OVARIAN GRANULOSA CELL
13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION W81757
VERSION W81757.1 GI:1392776
KEYWORDS EST
SOURCE house mouse.

```

```

51 GlyGlnAlaGArgGLeuAlaProAsnPhaArgTfPalalleProAsnAr 67
|||||
157 GGGCGAGGTGGCGAGTCCCTTAACTTCGATGGCCATTCGCCACNG 206
|||||
67 glnMetAsnAspGlyLeuGlyGlyAspGlyAspAspMetGluMetPheM 84
|||||
207 GCAGATGAATGACGGTGTGGGTGGAGATGGAGATGATATGGAAATGTCCA 256
|||||
84 eGlgLgLUmetArgGluIleArgGlyLysLeuArgGluLeuLeuArg 100
|||||
257 TGGAGAGATGAGAGAGATCCGGAGAACCTGATAGGCTACACGTGAGA 306
|||||
101 AsnAsyGLeuArgLeuMetGlyGluLeuSerAsnHisHsAspHisH 117
|||||
307 AATTGTCTACGATCCCTATGGGGAGCTGTCTAACCCACGACGATCACA 356
|||||
117 aAspGluPhuCysLeuMetPro 124
|||||
357 TGAATGAACTCTGCGCTTATGCT 378
|||||
seq_name: gb_est1:AW476468
seq_documentation: 628 bp mRNA
LOCUS AW476468
EST
24-FEB-2000
22007

```

DEFINITION
similar to NC1-CGAP_L033; MS MUSCULUS CLONE ONE
967602.YI NC1-CGAP_L033; MS MUSCULUS CLONE ONE
to G818188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HG874
sequence, mRNA sequence.
EST.
Accession Number AM474688.1 GI:7046574
Version AM474688
Keywords house mouse.
Source EST.
Organism Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NC1-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgap@rscg.nhlbi.nih.gov
Tissue Procurement Albert Smith, Ph.D.
Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
Condon Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Distribution: NC1-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium,
http://llnl.gov/dbfp/image/image.htm

```

K:      47
S:      :met:-40RP from Gibco
F:      quality sequence stop: 456.
       Location/Qualifiers
         1..628
           /organism="Mus musculus"
           /strain="C58CH II0090"
           /db_xref="IMAGE:2937027"
           /clone_lib="NCI CGAP Lu33"
           /tissue.type="pooled lung tumors"
           /lab_host="DH10B (phage-resistant)"
           /notes="Organ: lung; Vector: pT7D-Pac (Pharmacia) with a
modified polynucleotide linker. Site1: NotI; Site2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors. The template DNA was digested with NotI and
ligated into the modified pT7D-Pac vector. Library
size=1500bp. Sequence: ATGTTCTGGTGTATTTTTTTT 3'".
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```

motor

```

BASE COUNT      162 a  157 c  172 g  93 t  1 others
ORIGIN

alignment_scores:
  Quality: 688.00      Length: 124
  Ratio: 5.548        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 98.387

alignment_block:
US-09-327-750D-12 x BE334877 ..
Align seg 1/1 to: BE334877 from: 1 to: 585

```

Align seg 1/1 to: BE334877 from: 1 to: 585

1 MetAspValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAs 17
166 ATGGCCAAATGTCACCCAGGAAACAAAGAGATGGAGCCGCCCTGCAGAA 215
17 nGlyGluGluAspArgProValGlyGlyGlyGluGluHisGlnProAlaG 34
216 TGGACAGGAAGACGCCCTGTGGAGGAGGTGAGAGGCCACACGCTGCTG 265
34 LysAsnAsnAsnAsnAsnHisHisAsnHisAsnHisAsnHisArgArg 50
266 CAAC 315
51 GlyGlnAlaAspArgLeuAlaProAsnPheAspArgProAlaLeuProAsnArg 67
316 GGCCAGGCTGCCGACGATGGCCCTCACTTCCGATGGGCCTATCCCAACAG 365
67 GlnMetAsnAspGlyLeuGlyGlyAspGlyAspAspMetGluMetPheM 84
366 GCAGTGAATGTCGGTGTGGGTGGAGATGGAGATGATATGCAATATGCA 415
84 etGGluMetArgGluLeuArgArgGlyLeuAspGlyGluGluGluLeuArg 100
416 TGGAGAGATCAGAGAGATCCGGAAGACTTGGGAGGTACAGCTGAGA 465
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
466 AATGTGTCACCATCCCTTGTGGGAGAGGTGCTTAACCAACCAACACCA 515
117 GATGAATTCGTGCTTATGCT 537

seq_name: gb_est2:W64711	seq_documentation_block:	590 bp	mRNA	EST
	LOCUS	W64711		
	DEFINITION	gb47q01.r1 Soares mouse embryo NM013.5 14.5 Mus musculus cDNA		10-JUN-1995
	IMAGE	IMAGE:373872.5		
	DESCRIPTION	13.0 KD PROTEIN HGR74 (HUMAN); mRNA sequence.		

W64711.1 GI:1372353
EST.
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
(bases to 590)
Marz, S., Haller, T., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Hildebrand, M., LeMay, J., Martin, J., Morris, R.,
Schellenberg, K., Stepien, M., Tan, F., Unwin, K., Woodard,
Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tue

alignment_scores:

Quality:	688.00	Length:	124
Ratio:	5.548	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	98.387

alignment block:

117 sAspGluPheCysLeuMetPro 124

117 sAspGluPheCysLeuMetPro 124

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 15:42:51 ; Search time 144 Seconds
4167.553 Million cell updates/sec

Title: US-09-327-750D-28
Perfect score: 700
Sequence: 1 acgagcgtcgcgcagcagc.....gggctgtgttcagtgta 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	693.6	99.1	700	22	AAF23528
2	307.4	43.9	891	22	AAF23529
3	284.6	40.7	917	21	AAF21748
4	253.6	36.2	532	21	AAC01005
5	192.2	27.5	485	21	AAC01004
6	147.2	21.0	187	21	AAH87187
7	127.2	18.2	825	22	AAH83577
8	80.2	11.5	1239	22	AAH13750
9	80.2	11.5	1364	22	AAC85548
10	80.2	11.5	401	21	AAC01003
11	58	8.3	401	21	AAC01003

12	56.4	8.1	792	22	AAH45143
13	55.8	8.0	767	22	AAH75810
14	55.8	8.0	898	22	AAF59611
15	55.6	7.9	662	21	AAC03880
16	55.6	7.9	832	21	AAC03879
17	55.6	7.9	858	22	AAI03671
18	55.6	7.9	858	22	AAI03671
19	54.2	7.7	862	22	AAI58581
20	50.4	7.2	288	21	AAC06182
21	44.6	6.4	5552	21	AAAS8309
22	42.2	6.0	128	21	AAC12100
23	41.4	5.9	830	22	AAI22045
24	41.4	5.9	830	22	AAI47337
25	41.4	5.9	830	22	AAI07740
26	40.8	5.8	279	21	AAC06181
27	40.8	5.8	381	22	AAI13326
28	40.8	5.8	381	22	AAI13326
29	40.8	5.8	381	22	AAI02485
30	40.4	5.8	10732	21	AAAI0594
31	39.2	5.6	1635	22	AAI22780
32	39.2	5.6	1635	22	AAI48082
33	39.2	5.6	1635	22	AAI08454
34	39.2	5.6	1973	22	AAI13556
35	39.2	5.6	1973	22	AAI34918
36	39.2	5.6	1973	22	AAI03446
37	38.8	5.5	242	21	AAH03084
38	38.4	5.5	1155	21	AAH17742
39	38.4	5.5	1155	21	AAC38965
40	38.4	5.5	1155	21	AAC76371
41	38.4	5.5	1733	22	AAAS03059
42	37.6	5.4	439	22	AAI22718
43	37.6	5.4	439	22	AAI48016
44	37.6	5.4	439	22	AAI08383
45	37.6	5.4	1709	21	AAF15595

ALIGNMENTS

RESULT 1
ID AAF23528 standard: DNA: 700 BP.
XX AAF23528;
AC AAF23528;
DT 22-MAR-2001 (first entry)
XX Mouse NADE DNA.
XX Mouse NADE DNA.
KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappaB; ds.
OS Mus sp.
PN W020007578-A2.
XX 14-DEC-2000.
XX 07-JUN-1999; 99US-0327750.
XX 07-JUN-1999; 99US-0327750.
XX (LIVE) UNIV COLUMBIA NEW YORK.
XX Sato T.
XX WPI; 2001-061707/07.
XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
XX the gene encoding NADE, useful for modulating the activity of p75NTR
XX and for detecting neurodegenerative diseases.
XX Claim 12; Fig 1; 134pp; English.

Human brain expres
Human x chromosome
Human cell cycle a
Human secreted pro
Human secreted pro
Human secreted pro
Human polynucleoti
Human polynucleoti
Human secreted pro
pDIP/PyCSP.1 plasm
Human secreted pro
Probe #11978 for g
Probe #15023 used t
Probe #7731 used t
Human secreted pro
Probe #2512 used t
Probe #2476 used t
Gene encoding a su
Probe #12713 for g
Probe #16768 used
Probe #8445 used t
Probe #3489 for ge
Probe #3604 used t
Probe #3439 used t
Human cDNA sequenc
Arabidopsis thalia
Human ORF ORF1926
Human diagnostic a
Probe #12651 for g
Probe #16702 used t
Probe #8374 used t
Human prostate can

CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the protein-coding
CC the full-length cDNA without any specialized methods. AAH0166 to AAH1362 and
CC AAH1363 to AAH19742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

Query Match 11.5%; Score 80.2; DB 22; Length 1229;
Best Local Similarity 56.2%; Pred. No. 1.8e-12;
Matches 269; Conservative 0; Mismatches 153; Indels 57; Gaps 4;
QY 143 gccacgaatgacacaggaacaaatctcatgccaatgtccacacgagaacga 202
Db 196 gtccaaagggaactagcgcgaacattctcaacggggaatgcccaacagaacga 255
QY 203 agagctggagcag---ccctcgagaatggcaggaacacgcctctgtggaggaggtga 259
Db 256 agagggggggcagggcccccacgaggaatgagaagaatcccctctgtggaggaggtga 315
QY 260 ggccacacagcctgctgcacacacacacacacacacacacacacacacacac 319
Db 316 agggcagagcctggagggaatctca-----
QY 320 ccgaaggagcagcctgcgcgactgcctcaacttcgactggccattcccaacaggca 379
Db 342 -ggcggggggcagattagggactgttcccaatttctgagggcattacctaaggca 399
QY 380 gatgaatgcgggtgggtggagatggagatgataatggaatgttcacgagagatgag 439
Db 400 tatgtgacaa-----tgagcgagagagatgataatggaatgttcacgagagatgag 453
QY 440 agagctcccgagaaagcctggagagcagcctgcagaaattgtctacgactcattggtg 499
Db 454 ggaatcaagagaagatgagggaacagcagatgaggcactatgctgcttccaaa---- 509
QY 500 gtagctgtcaac 559
Db 510 -----ctctgaacctgacacacacacacacacacacacacacacacacacac 561
QY 560 attccctgagatcattatgactgcctgcctgtagccttccctccctgactttct 618
Db 562 agtttctgaggttaattgaaactgctttacaaagcttgaattttgtgatttact 620

RESULT 10
AAC85548
ID AAC85548 standard; cDNA; 1364 BP.

XX AAC85548;

XX 04-JUN-2001 (first entry)

DE cDNA encoding CDIFF-4, Incyte ID No. 1990956CB1.

KW Human; cell differentiation; CDIFF: agonist; antagonist; epilepsy;
KW cell proliferation; Alzheimer's disease; schizophrenia disorder;
KW arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 250..612
XX /tag= a
XX /product= "CDIFF-4"

XX WO200119860-A2.

XX 22-MAR-2001.
XX 14-SEP-2000; 2000WO-US25435.
XX 15-SEP-1999; 99US-0154140.
XX 06-DEC-1999; 99US-0169155.
XX (INCY-) INCYTE GENOMICS INC.
XX Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;
PI Baughn MR, Lu DAM, Bandman O, Shih LL, Patterson C;
XX HPJ: 2001-211447/21.
XX P-FSDB: AA847126.
XX Isolated polypeptides and polynucleotides involved in cell
PT differentiation are used for treatment, prevention and diagnosis of
PT cell proliferative, developmental and neurological disorders e.g.
PT cancer and Alzheimer's disease.
XX Claim 5; Page 121; 137pp; English.
XX The sequences given in AAC85548-72 encode human polypeptides involved
CC in cell differentiation (CDIFF). CDIFF polypeptides and agonists of
CC these are used to treat a disease or condition associated with
CC decreased expression of functional CDIFF. An antagonist of CDIFF is
CC used to treat a disease or condition associated with over expression
CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,
CC prevention and diagnosis of cell proliferative, developmental and
CC neurological disorders, such as Alzheimer's disease, schizophrenia
CC and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus
CC and REX-3. This sequence maps to chromosome 1 within the interval from
CC 137.4 to 138.0 centimorgans, and to the X chromosome within the interval
CC from 104.9 to 150.3 centimorgans.
XX Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;

Query Match 11.5%; Score 80.2; DB 22; Length 1364;
Best Local Similarity 56.2%; Pred. No. 1.8e-12;
Matches 269; Conservative 0; Mismatches 153; Indels 57; Gaps 4;
QY 143 gccacgaatgacacaggaacaaatctcatgccaatgtccacacgagaacga 202
Db 255 gtccaaagggaactagcgcgaacattctcaacggggaatgcccaacagaacga 314
QY 203 agagctggagcag---ccctcgagaatggcaggaacacgcctctgtggaggaggtga 259
Db 315 aggaggggagcagggcccccacgacgaatgagaagaatcccgcttggaggagggtga 374
QY 260 ggccacacagcctgctgcaacacacacacacacacacacacacacacacacac 319
Db 370 agggcagaagcctggagggaatctca-----
QY 320 ccgaaggagcagcctgcgcgactgcctcaacttcgactggccattcccaacaggca 379
Db 401 --ggcggggggagttaggcgactgtccctaatcttcgattggccattacgaagca 458
QY 380 gatgaatgacgggtgggtggagatggagatgataatggaatgttcacgagagatgag 439
Db 459 tattgagcaca-----tgagcgagagagatgataatggaatgttcacgagatgat 512
QY 440 agagatcccggaagagcttagggagctacagctgagaaattgtctacgactcctatggg 499
Db 513 ggaatacgaagaagatgagggaacagcagatggcactatgctgtcccaaa---- 568
QY 500 gtagctgtcaac 559
Db 569 -----ctctgaacctgacacacacacacacacacacacacacacacacacac 620

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 16:36:39 ; Search time 144 Seconds
(without alignments)
5304.700 Million cell updates/sec

Title: US-09-327-750D-29
Sequence: 891
1 accccatccccctctat.....ataaagcaattaaagc 891

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N_Geneseq_1101.*
- 1: /SIDS2/gcgdata/geneseq/NA1980.DAT.*
 - 2: /SIDS2/gcgdata/geneseq/NA1981.DAT.*
 - 3: /SIDS2/gcgdata/geneseq/NA1982.DAT.*
 - 4: /SIDS2/gcgdata/geneseq/NA1983.DAT.*
 - 5: /SIDS2/gcgdata/geneseq/NA1984.DAT.*
 - 6: /SIDS2/gcgdata/geneseq/NA1985.DAT.*
 - 7: /SIDS2/gcgdata/geneseq/NA1986.DAT.*
 - 8: /SIDS2/gcgdata/geneseq/NA1987.DAT.*
 - 9: /SIDS2/gcgdata/geneseq/NA1988.DAT.*
 - 10: /SIDS2/gcgdata/geneseq/NA1989.DAT.*
 - 11: /SIDS2/gcgdata/geneseq/NA1990.DAT.*
 - 12: /SIDS2/gcgdata/geneseq/NA1991.DAT.*
 - 13: /SIDS2/gcgdata/geneseq/NA1992.DAT.*
 - 14: /SIDS2/gcgdata/geneseq/NA1993.DAT.*
 - 15: /SIDS2/gcgdata/geneseq/NA1994.DAT.*
 - 16: /SIDS2/gcgdata/geneseq/NA1995.DAT.*
 - 17: /SIDS2/gcgdata/geneseq/NA1996.DAT.*
 - 18: /SIDS2/gcgdata/geneseq/NA1997.DAT.*
 - 19: /SIDS2/gcgdata/geneseq/NA1998.DAT.*
 - 20: /SIDS2/gcgdata/geneseq/NA1999.DAT.*
 - 21: /SIDS2/gcgdata/geneseq/NA2000.DAT.*
 - 22: /SIDS2/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	889.4	99.8	891	22	AAF23229
2	564.4	63.3	917	21	AAF21748
3	310.4	38.5	724	21	AAF21528
4	298.8	33.5	485	21	AAF21004
5	273.8	30.7	426	22	AAF23586
6	273.8	30.7	426	22	AAF23586
7	160.8	18.0	169	16	AAT21034
8	150.8	16.9	187	21	AAH87147
9	141.2	15.8	865	22	AAH03517
10	141.2	15.8	1229	22	AAH13750
11	141.2	15.8	1364	22	AAC85548

12	108.2	12.1	792	22	AAH45143
13	104	11.7	998	22	AAF59611
14	102.4	11.5	767	22	AAH75810
15	102.4	11.5	767	22	AAH75810
16	91.4	10.7	858	22	AAH75810
17	91.4	10.7	858	22	AAH75810
18	89.8	10.1	662	21	AAH03880
19	89.8	10.1	662	21	AAH03880
20	81	9.1	698	21	AAH03879
21	81	9.1	936	22	AAF58252
22	81	9.1	936	22	AAF58252
23	81	9.1	936	22	AAF58252
24	81	9.1	936	22	AAF58252
25	81	9.1	936	22	AAF58252
26	74.6	8.4	1197	21	AAH72409
27	69.2	7.7	936	22	AAF58252
28	67.4	7.6	936	22	AAF58252
29	67.4	7.6	936	22	AAF58252
30	67.4	7.6	936	22	AAF58252
31	67.4	7.6	936	22	AAF58252
32	67.4	7.6	936	22	AAF58252
33	67.4	7.6	936	22	AAF58252
34	66.4	7.5	128	21	AAH12100
35	64.8	7.3	397	22	AAH14082
36	64.8	7.3	397	22	AAH14082
37	64.8	7.3	397	22	AAH14082
38	58.6	6.6	1170	22	AAH69527
39	58.6	6.6	1170	22	AAH69527
40	56.4	6.3	1521	22	AAH159746
41	56.4	6.3	1521	22	AAH159746
42	55.4	6.2	1309	21	AAH30950
43	52.8	5.9	285	22	AAH21436
44	52.8	5.9	285	22	AAH21436
45	52.8	5.9	285	22	AAH21436

ALIGNMENTS

RESULT	1
XX	AAF23529 standard; DNA; 891 BP.
XX	AAF23529;
AC	AAF23529;
XX	22-MAR-2001 (first entry)
DE	Human NADP DNA.
XX	Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW	Neurogenetic disease; NF-kappaB; ds.
XX	Homo sapiens.
XX	WO200075278-A2.
PN	14-DEC-2000.
PD	07-JUN-2000; 2000WO-US15621.
PF	07-JUN-1999; 99US-0327750.
PR	(UYCO) UNIV COLUMBIA NEW YORK.
XX	Sato T.
XX	WPI; 2001-061707/07.
XX	New p75-neurotrophin receptor-associated cell death executor (NADE) and
PT	the gene encoding NADE, useful for modulating the activity of p75NTR
PT	and for detecting neurodegenerative diseases .
XX	Disclosure; Fig 1; 134pp; English.

QY 343 tggagcag---cctatgcagaatggagagagacccgcttctggagggagtggaagcc 399
Db 261 gggagcagcccccacgcagatgaagagaaaccccgccattctggaggggaggaagcc 320
QY 400 accgctcgcaggaat---cgacgggacagcgcgcgcgacttgcctccataatttcgat 456
Db 321 agagcctggaggaatacatcagcgggggagtgagggagctgcctccataatttcgat 380
QY 457 gggccatccacacagatcagatgaatggatggatggatggatggatggatggatggat 516
Db 381 gggccatccatcattagcagat---tgagcacaatgaagcagagagatgagtaganaa 434
QY 517 tattctatggagagatagagaaatcagagaaacttggagcgcagtgaggaatt 576
Db 435 ggttctggggcagatagagaaatcagagaaacttggagcgcagtgaggaatt 494
QY 577 gctcgatatacttctgggggagctcttcttaccatgacacatcagtgaggaatt 636
Db 495 atagcgcttccaaa-----ctctggaactgacacattatgactttgccc 542
QY 637 ttatgcctgactcctcatttacc 662
Db 543 tcatacctgaaatcctaaaagtcttc 568

RESULT 10
ID AAH13750 standard; cDNA; 1229 BP.
XX AAH13750;
XX AAH13750;
DT 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:10656.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EPI074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX MPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX Claim 8; SEQ ID 10656; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC identification of a disease. The primers allow obtaining of the full-length
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

QY Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;
Best Local Similarity 15.8%; Score 141.2; DB 22; Length 1229;
Matches 264; Conservative 0; Mismatches 98; Indels 24; Gaps 4;
QY 283 aaaaacacacagaaaaaaatctctatcattggaataattccaccaggaagaaacgaagaga 342
Db 201 agaggaactagcggcaaacatctcaacggggggaatgcccaacagaaacgaagagag 260
QY 343 tggagcag---cctatgcagaatggagagagacccgcttctggagggagtggaagcc 399
Db 261 gggagcagcccccacgcagatgaagagaaaccccgccattctggaggggaggaagcc 320
QY 400 accgctcgcaggaat---cgacgggacagcgcgcgcgacttgcctccataatttcgat 456
Db 321 agagcctggaggaatacatcagcgggggagtgagggagctgcctccataatttcgat 380
QY 457 gggccatccacacagatcagatgaatggatggatggatggatggatggatggatggat 516
Db 381 gggccatccatcattagcagat---tgagcacaatgaagcagagagatgagtaganaa 434
QY 517 tattctatggagagatagagaaatcagagaaacttggagcgcagtgaggaatt 576
Db 435 ggttctggggcagatagagaaatcagagaaacttggagcgcagtgaggaatt 494
QY 577 gctcgatatacttctgggggagctcttcttaccatgacacatcagtgaggaatt 636
Db 495 atagcgcttccaaa-----ctctggaactgacacattatgactttgccc 542
QY 637 ttatgcctgactcctcatttacc 662
Db 543 tcatacctgaaatcctaaaagtcttc 568

RESULT 11
ID AAC85548
XX AAC85548 standard; cDNA; 1364 BP.
XX AAC85548;
XX AAC85548;
DT 04-JUN-2001 (first entry)
XX cDNA encoding CDIFF-4, Incyte ID NO. 1990956CB1.
XX Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;
XX cell proliferation; Alzheimer's disease; schizophrenia disorder;
XX arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 250..612
FT CDS /tag= a
FT /product= "CDIFF-4"
XX


```
CC specification.
XX Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;
SQ

Query Match      11.5%; Score 102.4; DB 22; Length 862;
Best Local Similarity 55.3%; Pred. No. 1.5e-20;
Matches 345; Conservative 0; Mismatches 241; Indels 38; Gaps 6;

Qy 283 aaaaaaacacccgaaaaaaataatctcatctgacaaatattccacgggaaacgaagaga 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 aagaggacgagcgttaaacatctcatctcgaggaaatgcaacccaggaaaatgatgaaa 295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 343 tggagagccctatgagaatgagaggaaga-----ccgcctttggagagag 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 aagatgaaggagcgaagtgtctaagaaggagcccttggccctacccttgaatgta 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 391 gtgaaggccacacgacctgcaggaagaa---tcagcgggagacgcgcgcgacttgcacctta 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 gtgaatactgtgctcctagaggaacgcgtagcgggtccgcgttaggcagcccatctgc 415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 448 atttcgatggccatcccataggcagatcaatgatggatgggtggagatggagatg 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 416 agtatgaggggacataatgcataggtcttgagagccacacggaagatgagagagaga 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 508 atatggaaaatattcatggagagatgagagaatcagaagaataacttaaggagctgcagt 567
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 atatgaaaggattggggagggtgagacagctgatggaaaagctgagggaaaagcagt 535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 568 tdaggaaattgtcgtatctctatggggagctctctaatcaccatgacctcatgatg 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 536 tdagtcatagttcggggcagtcagacactgaccc---tcaccatgacctcagatg 592
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 628 aatttgcctatgacctgacctaccatttatatgagatgaatgaatgaattatctccgc 687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 593 agtttgcctatgacctgacctgattgttctcctgaagttaattagggagaccctgc 652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 688 tgtttctttctctgatttctcctaataatgaccttactgatccgtttgtgtgaacc 747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 653 ttcctaaacttaca-----catttgtgtgaaccttgcgtaaaacy 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 748 tatgtattccatggtcaagtggtctgtgtgcaagcttctctatttgaagatgctt 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 695 tttagagttaccatattcttgggtccctcattaccagttcttaaaagatgtgttt 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 808 ttg-cactcaagtgaagtctctgcagcagtag-ttccaccatttgcatggaanaattt 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 755 ttgaccagtttgtlaagtctctgcagcagagaggtttaccctattgcatggaagatgc 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 866 aaagccataaagaacatttaanaa 889
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 815 tcattatatttgaaatcaata 838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 11, 2002, 16:36:43
Job time: 3232 sec

OM of: US-09-327-750D-35 to: EST:* out_format : pfs

Date: Mar 11, 2002 2:16 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=frame+2n.model -DEV=xlh
-q/cgn2.1/USPTO.spool/US09327750/runat_l1032002_l01153_20308/app_query.fasta_1.1472
-DB=EST -OEMT=FASTAP -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.500
-MINMATCH=0.100 -LOOPECI=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750.ecgn1.1.5654
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-327-750D-35

Query length: 117

Database: EST:

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 4085.940000

score_list:

Sequence	Strd Orig	Zscore	EScore	Len	Documentation
gb_hnc:AK010400	+	611.50	1049.04	789	AK010400 Mus musculus ES cells
gb_est2:BG083261	+	611.50	1049.02	792	BG083261 H3086C08-5 NIA Mouse 1
gb_est1:AA473525	+	605.00	1041.70	458	AA473525 v378a09.r1 Barstead MF
gb_est2:BG070341	+	575.50	987.47	747	BG070341 H3086C08-3 NIA Mouse 1
gb_est2:BF723075	+	568.50	977.99	453	BF723075 mab27e12.r1 Soares NHE
gb_est1:BE323774	+	544.50	937.49	469	BE323774 us27h02.y1 Soares NHE
gb_est1:W17712	+	541.00	934.14	317	W17712 mb77e04.r1 Soares NHE
gb_est1:A1272867	+	539.50	927.44	568	A1272867 EST24582 Normalized 1
gb_est2:RI289546	+	511.50	878.38	924	RI289546 UI-R-DKO-clf-c-12-0-01
gb_est1:AA117439	+	508.50	869.71	312	AA117439 m21e12.r1 Beddington
gb_est2:W35893	+	484.50	834.09	472	W35893 mc53g09.r1 Soares mouse
gb_est2:BF722027	+	479.00	828.16	286	BF722027 mab27e12.x1 Soares NHE
gb_est2:W54487	+	469.50	808.38	465	W54487 md09c11.r1 Soares mouse
gb_est1:AL118340	+	463.50	798.74	432	AL118340 V911b26 Beddington mc
gb_est1:AW251884	+	459.00	789.25	540	AW251884 UI-R-BJO-adn-b-06-0-01
gb_est1:AA104150	+	426.50	734.87	433	AA104150 mc50h03.r1 Life Tech
gb_est2:W85403	+	426.50	734.73	442	W85403 mf58d02.r1 Soares mouse
gb_est1:AI012535	+	373.00	642.62	444	AI012535 EST206986 Normalized 1
gb_est1:AV137854	+	368.00	636.78	300	AV137854 AV137854 Mus musculus
gb_est1:AA516739	+	365.00	631.85	290	AA516739 v8h4c04.r1 Knowles Sol
gb_est2:W61422	+	363.50	630.70	237	W61422 md92g02.r1 Soares mouse
gb_est1:AL517301	+	355.50	609.93	639	AL517301 AL517301 LTI.NFL011.NE
gb_est2:BE883414	+	355.50	609.07	722	BE883414 601509167F1 NIH.MGC.71
gb_est1:AL544124	+	355.50	607.95	845	AL544124 AL544124 LTI.NFL006.PI
gb_est1:AU117899	+	355.50	607.79	865	AU117899 AU117899 HEMBA1 Homo S
gb_est2:BG122197	+	355.50	607.78	866	BG122197 602349748F1 NIH.MGC.90
gb_est1:AL545120	+	355.50	607.47	905	AL545120 AL545120 LTI.NFL006.PI
gb_est2:BG709427	+	355.50	607.29	929	BG709427 602673481F1 NIH.MGC.96
gb_est2:BG035675	+	355.50	606.94	975	BG035675 602325745F1 NIH.MGC.96
gb_est1:AL560850	+	355.50	606.89	983	AL560850 AL560850 LTI.NFL010.BC
gb_est1:AL539004	+	355.50	606.61	1023	AL539004 AL539004 LTI.NFL013.BC
gb_est1:AL549875	+	355.50	606.46	1044	AL549875 AL549875 LTI.NFL005.F
gb_est2:BF930215	+	355.50	602.88	579	BF930215 IL5-NT0227-111200-319
gb_est1:AL534630	+	350.50	602.18	566	AL534630 AL534630 LTI.FL013.FB
gb_est1:AL534630	+	349.00	595.28	1043	AL534630 LTI.FL013.FB
gb_est1:AW385267	+	347.50	597.43	534	AW385267 RCO-UT0001-261199-011
gb_est2:BG527431	+	347.50	594.82	772	BG527431 602557227F1 NIH.MGC.59
gb_est2:BG009563	+	346.50	595.11	581	BG009563 QV1-GN0319-021200-526
gb_est1:AL582060	+	346.50	592.16	1037	AL582060 AL582060 LTI.NFL010.B
gb_est2:BI040887	+	344.50	592.88	489	BI040887 QV3-NT0278-120201-502

gb_est1:AL573904 - 344.50 586.55 1.6e-23 1182 | AL573904 AL573904 LTI.NFL010

gb_est2:BI040889 - 340.50 584.98 2.0e-23 565 | BI040889 QV3-NT0278-120201-

gb_est1:AW934883 + 338.50 580.35 3.7e-23 668 | AW934883 RCL-DT0001-230200-

gb_est1:AW934905 + 338.50 580.18 3.8e-23 685 | AW934905 RCL-DT0001-230200-

gb_est1:AW934769 + 338.00 579.56 4.1e-23 662 | AW934769 RCL-DT0001-230200-

seq_name: gb_hnc:AK010400

seq documentation block:

LOCUS AK010400 789 bp mRNA HTC 05-JUL-2001

DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library, clone:2410004M13, full insert sequence.

ACCESSION AK010400

VERSION AK010400.1 GI:12845816

KEYWORDS CAP Trapper.

SOURCE clone:lib:RIKEN full-length enriched mouse cDNA library

clone:2410004M13.

Mus musculus

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 789)

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome research. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3 (bases 1 to 789)

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer

JOURNAL Genome research. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4 (bases 1 to 789)

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5 (bases 1 to 789)

AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Hanakawa,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Inotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koyama,K., Numazaki,R., Oono,M., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Sakai,C., Sakai,K., Sano,H., Sakai,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,M., Takahashi,F., Tanaka,T., Teijima,Y., Toyota,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-resescg.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

This clone is available royalty-free through LINL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:511536
 Seq primer: -28ml3 rev2 Et from Amersham
 High quality sequence stop: 417.

FEATURES
SOURCE

```

1. 458
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:872056"
/clone_lib="Barstead MPLRBI"
/sex="mixed"
/tissue_type="Kidney"
/dev_stage="5 weeks"
/lab_host="DH10B"

```

```
/note=Vector: pT730-Pac (Pharmacia) with a modified  
polylinker. Site 1: EcoRI; Site 2: NotI; 3rd strand cDNA  
was primed with a Not I- oligo(dn) primer [5,  
TGTTACGAATCGAAGTAGGAGCGCCGCCGGCTTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
[CATGATTCGTGATC], digested with Not I and cloned into the  
Not I and Eco RI sites of the modified pT7m3 vector.  
Library constructed by Bob Harstead.
```

BASE COUNT	147 a	88 c	115 g	108 t
ORIGIN	Library constructed by Bob			

alignment_scores:		
Quality:	605.00	Length: 116
Ratio:	5.261	Gaps: 0
Percent Similarity:	99.138	Percent Identity: 97.414

alignment_block:

US-09-327-750D-35 x AA473525

Align seg 1/1 to: AA473525 from: 1 to: 458

2 AlaSerLysLysGlnValIleuAspLeuThrValGluLysAspLysLy 117
 13 GCCAAATTTAAACAAGTCATCTGGGATCTCATCTGGGAAGAACA 62
 18 sasplysargcylglyLysAlaSerLysGlnSerGluGluProHis 35
 63 AGACAAAANAAGGTGGGAAGCCTCCAAAACAAAGTGAAGAAGACC 112
 35 IsLeuGluGluValGluAsnLysLysProGlyGlyAsnValArgLys 51
 113 ATCTGGAGAGGTTGAAAACAGAAGCCTGGGGGAATATGCCGAAGA 162
 52 ValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnArgHisVa 68
 163 GTCAGGCGCATCTGGCGCTAACTTCTCTGGGCCATACCAATAGGCATGT 212
 68 IAsPArgAsnGluGlyGlyAspValGlyArgPheValValGlnGlyT 85
 213 TGATCGCAATGAAGGGGAGAGGATGTTGGGAGATTCTAGTCAGGGA 262
 85 hrGluValLysArgLysThrThrGluGlnValArgProTyrArgArg 101
 263 CAGAAGTCAAGAAAGACTACGGAGCAGCAGGTCAGGCCTTACAGCGCT 312
 102 PheArgThrProGluProAspAsnHisTyrAspPheCysLeuIlePro 117
 313 TTCGGAACCCGGAACCTGACAATCATTCACACATTTTGGCTCATACCT 360

seq_name: qb_est2:BG070341

seq documentation block:

<hr/>					
seq_documentation_block:					
LOCUS	BG070341	747 bp	mRNA	EST	26-JAN-2001
DEFINITION	H3086C08-3 NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone				
DEFINITION	H3086C08 3', mRNA sequence.				
ACCESSION	BG070341				

alignment_scores:		
Quality:	611.50	Length: 118
Ratio:	5.226	Gaps: 1
Percent Similarity:	99.153	Percent Identity: 98.305

alignment_block:

US-09-327-750D-35 x BG083261

Align seg 1/1 to: BG083261 from: 1 to: 792

[illegible]

50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
 |||
 340 AGGAAGTCAGCGGACCTTGTCCTAACTTCTCTGGGCCATACCAAAATAG 389

66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
 |||||
 390 GCATGTTGATCGCAATGAAGGGGGAGAGGATGTTGGGAGATTGTGATGTC 439

83 lnGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyr 99
|||||
440 AGGGAACAGAGAAGTCAAGAGAAAGACTACGGAGCAGCAGGTGAGGCCTTAC 489

100 ArgArgPheArgThrProGluProAspAsnHistyrAspPheCysLeu116
|||||
490 AGCGGTTTCGGAACCCGGACCTGACAATCATTCAGCACTTTGCCTCAT 539

116 ePro 117
540 ACCT 543

seq_name: qb_est1:AA473525

seq_documentation_block:
LOCUS AA473525 458 bp mRNA EST
DEFINITION v978a09.r1 Barstead MPR1B1 Mus musculus CDNA clone IMAGE:872056 5'
similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA CELL 13.0 KD
PROTEIN HGR74.1; mRNA sequence.

PROTEIN HQK74. // mRNA sequence:
ACCESSION AA473525

AA473525.1 GI:2201752

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
1 (bases 1 to 458)
Marra, M., Hiller, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gessel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Thelings, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT
Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park 1

Tel: 314 286 1800
Fax: 314 395 1810

Fax: 314 286 1810
Email: mougeest@vnet.net

VERSION
KEYWORDS
SOURCE
ORGANISM

BC070341.1 GI:1252910
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 747)
Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
, T.S., Carter, M.G. and Ko, M.S.H.
Verification and initial annotation of NIA mouse 15K cDNA clone set
Unpublished (2001)
Other ESTs: H3086C08-5
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://igun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3086 row: C column: 08
Seq primer: -21M13 Forward
High quality sequence stop: 747
POLYA=yes.

FEATURES
source

1..747
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="H3086C08"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dt)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
Mol Genet 7: 1967-1978."
BASE COUNT 178 a 198 c 140 g 231 t
ORIGIN

alignment_scores:

Quality: 575.50 Length: 118
Ratio: 5.004 Gaps: 1
Percent Similarity: 97.458 Percent Identity: 96.610

alignment_block:

US-09-327-750d-35 x BC070341/rev ..
Align seg 1/1 to reverse of: BC070341 from: 1 to: 747
1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValClnLysAs 16
|||||
581 ATGGGATCCAAATTAACAAAGCTACTGATCTCCTGTCGGAAGA 532
|||||
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluP 33
|||||

531 CAAAAAGACAAAAAGGTGGGAGGCCTCCAAACAAGTGAAGAGAAC 482
33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
481 CCCACCATCTGGAAGAGGTGAACAAAGAGAGCCTGGGGAAATGTCGA 432
50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
|||||
431 AGGAAGTACAGGACCTTGTGCTTACCTTCTCTGGCCATACCAATAG 382
66 gHisValAspArgAsnGluGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 83
381 GCATGTTGATCGCAATGAAGGGGGAGAGCATGTTGGGAGATTGTAGTC 332
83 InGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyr 99
|||||
331 AGGAACAGAGTCTACAGAGAAAGACTACGGAGCAGCAGGTGAGCGT.TAC 283
100 ArgArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeu 116
|||||
282 AGCGGTTTCGGAACCCCGAGACCTGACAATCATTACGACITTTGCCTCAT 233
116 ePro 117
|||||
232 ACCT 229
seq_name: gb_est2:BF723075
seq_documentation_block:
LOCUS BF723075 519 bp mRNA EST 03-JAN-2001
DEFINITION mab27e12.y1 Soares_NMEBA_branchial_arch Mus musculus cDNA clone
IMAGE:3971447 5' similar to TR:Q9Y516 Q9Y516 D63519.2.3 ;, mRNA
sequence.
ACCESSION BF723075
VERSION BF723075.1 GI:12024077
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 519)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: mab27e12.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MG1:1471479

Seq primer: -40RP from Gibco
High quality sequence stop: 472.

FEATURES
source

Location/Qualifiers
1..519
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3971447"
/clone_lib="Soares_NMEBA_branchial_arch"
/tissue_type="branchial arches"
/dev_stage="embryo, 10.5 dpc"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA
was primed with a Not I - oligo(dt) primer [5',
TCTTACCAATCTGAAGTGGCGCGCATGATTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo.

BASE COUNT 158 a 109 c 170 g 81 t 1 others
ORIGIN

alignment_scores:
Quality: 568.50 Length: 113
Ratio: 5.168 Gaps: 1
Percent Similarity: 97.345 Percent Identity: 96.460

alignment_block:

US-09-327-750D-35 x BF723075 ..

Align seg 1/1 to: BF723075 from: 1 to: 519

```

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
181 ATGGCATCCAAATTAACCAAGTCATCTGATCTCAGTGTGGAGAAAGA 230
|||||
16 physLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluP 33
|||||
231 CAAAAAGACAAAAAGGTGGAGGCTCCAAACAAAGTGAAGAGAAC 280
|||||
33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
281 CCCACATCTCGAGAGAGTTGAAACCAAGAGCGCTGGGGAAATGTCCGA 330
|||||
50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
|||||
331 AGGAAAGTCAGCGGACTGTGCTTAACCTTCTGCGCCATACCTAATAG 380
|||||
66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
|||||
381 GCATCTTCATCGCAATGAAGGGGAGAGAGATGTTGGAGATTGTAGTGC 430
|||||
83 InGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyr 99
|||||
431 AGGGAACAGAGNGTCAAGAGAAAGACTACGAGCAGCAGGTGAGGCGCTTAC 480
|||||
100 ArgArgPheArgThrProGluProAspAsnHisTyrAsp 112
|||||
481 AGGCGTTCCTCCAGACCGCGGACCTGACAAATCATTACGAC 519
|||||
seq_name: gb_est1:BE333774

```

seq_documentation_block: 463 bp mRNA 14-JUL-2000
LOCUS BE333774
DEFINITION us27h06.y1 Soares_NMEBA_branchial_arch Mus musculus cDNA clone
IMAGE:3168347 5', similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA
CELL 13.0 KD PROTEIN HGR74. ;, mRNA sequence.
ACCESSION BE333774
VERSION BE333774.1 GI:9207550
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 463)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

image.llnl.gov/image/html/iresources.shtml

MGI:1063807

Seq primer: -40RP from Gibco.

FEATURES
source

```

Location/Qualifiers
1..463
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="3168347"
/clone_lib="Soares_NMEBA_branchial_arch"
/tissue_type="branchial arches"
/lab_host="embryo, 10.5 dpc"
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TGTTACCAATCTGAAGTGGAGCGCGCATGATTTTGTGTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not  
I and Eco RI sites of the modified pT73 vector. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

```

BASE COUNT 155 a 91 c 147 g 70 t
ORIGIN

alignment_scores:
Quality: 544.50 Length: 113
Ratio: 4.950 Gaps: 1
Percent Similarity: 97.345 Percent Identity: 95.575

alignment_block:

US-09-327-750D-35 x BE333774 ..

Align seg 1/1 to: BE333774 from: 1 to: 463

```

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
126 ATGGCATCCAAATTAACCAAGTCATCTGATCTCAGTGTGGAGAAAGA 175
|||||
16 physLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluP 33
|||||
176 CAAAAAGACAAAAAGGTGGAGGCTCCAAACAAAGTGAAGAGAAC 225
|||||
33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
226 CCCACATCTCGAGAGAGTTGAAACCAAGAGCGCTGGGGAAATGTCCGA 275
|||||
50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
|||||
276 AGGGAACAGAGNGTCAAGAGAAAGACTACGAGCAGCAGGTGAGGCGCTTAC 325
|||||
66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
|||||
326 GCATCTTCATCGCAATGAAGGGGAGAGATCTTGGAGATTGAGTGC 374
|||||
83 InGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyr 99
|||||
375 AGGGAACAGAGAGTCAAGAGAAAGACTACGAGCAGCAGGTGAGGCGCTTAC 424
|||||
100 ArgArgPheArgThrProGluProAspAsnHisTyrAsp 112
|||||
425 AAGCGTTCCTCCAGACCGCGGACCTGACAAATCATTACGAC 463
|||||
seq_name: gb_est2:W17712

```

seq_documentation_block:

LOCUS W17712
DEFINITION mb77e04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:335454 5', similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA
CELL 13.0 KD PROTEIN HGR74. [1] ;, mRNA sequence.
ACCESSION W17712
VERSION W17712.1 GI:12992113

KEYWORDS SOURCE ORGANISM

EST.
house mouse
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 317)
AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucab, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE
The WashU-HMI Mouse EST Project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:216854
Seq primer: mob.REGA+ET
High quality sequence stop: 291.

FEATURES source

Location/Qualifiers
1..317
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:335454"
/dev_stage="19.5 dpc total fetus"
/lab_host="Dhl0B (ampicillin resistant)"
/note="vector: pT73D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCGAAGTGGGAGCGCGCATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."
BASE COUNT 108 a 63 c 87 g 59 t
ORIGIN

alignment_scores:

Quality: 541.00 Length: 101
Ratio: 5.410 Gaps: 0
Percent Similarity: 99.010 Percent Identity: 98.020

alignment_block:

US-09-327-750D-35 x W17712 ..

Align seg 1/1 to: W17712 from: 1 to: 317

17 LysLysAspLysArgGlyGlyAlaSerLysLysInserGluGluGluPr 33
1 AAAAAAGACAAAAAGGTGGGAGGCTCCAAACAAGTGAAGAGAACCC 50
33 oHsHisLeuGluValGluAsnLysLysProGlyGlyAsnValArg 50
51 CCACCATCTGGAAGAGGTGGAACAAGAGAGGCTGGGGAAATGTCGGA 100
50 rGlyValArgArgLeuValProAsnPhelLeuTirPalalileProAsnArg 66
101 CGAAAGTCAGCGAGCTTGTGCTTAACCTTCTCTGGGCCCATACCAATAGG 150
67 HisValAspArgAsnGluGlyGlyAlaAspValGlyArgPheValValG 83
151 CATGTTGATCCCAATGATGAGGGGAGGAGATGTTGGGAGATTGTTAGTCA 200

83 nGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyrA 100
111
201 GGGAAACAGAGTCAAGAGAAAGACTACGGAGCAGCAGGTGAGGCTTACA 250
100 rGArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeuIle 116
111
251 GGCCTTTCGAAACCCCGGAACCTGCATCATCATTAGACTTTTGCTCAT 300
117 Pro 117
111
301 CCT 303

seq_name: gb_estl:AI227867

seq_documentation_block: 568 bp mRNA EST 20-JAN-1999

LOGUS AI227867
DEFINITION EST224562 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
RBCRM82, 3' end, mRNA sequence.

ACCESSION AI227867
VERSION AI227867.1 GI:3811754
KEYWORDS EST.

SOURCE Rattus sp.

ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 568)

AUTHORS Lee, N.H., Glodex, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R. and Adams, M.D.

TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat

JOURNAL Gene Index

COMMENT Unpublished (1998)

Other ESTs: TC58086

Contact: Lee, NH

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529

Fax: (301)-838-0208

Email: nhlee@tigr.org

Seq primer: M13-21.

FEATURES

Location/Qualifiers

1..568

/organism="Rattus sp."

/db_xref="ATCC (inhost):2036891"

/db_xref="taxon:10118"

/clone="RBCRM82"

/note="lib="Normalized rat brain, Bento Soares"

/note="Organ: brain; Vector: pT73pac; Site_1: EcoRI;

Site_2: NotI"

BASE COUNT 140 a 143 c 108 g 177 t

ORIGIN

alignment_scores:

Quality: 539.50 Length: 118

Ratio: 4.860 Gaps: 1

Percent Similarity: 94.068 Percent Identity: 87.288

alignment_block:

US-09-327-750D-35 x AI227867/rev ..

Align seg 1/1 to reverse of: AI227867 from: 1 to: 568

1 MetalaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16

528 ATGGCGTCCAAAGTCAACAAGTCATACCTGGATCTCACTGTGGAGAAAGA 479

16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGlu 33

478 CAAGAAACAAACAAAAGGTGGGAGGCTCCAAACAAAGTGAAGAAGAT 429

33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49

```

|||||
428 CCACATCTGGAGAGGTTGAACACAGAGCCTGGGGCAATGTCAGG 379
|||||
50 ArgLysValArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
|||||
378 AGGAAGTCAAGCGAGCTTGCTGAACCTTCTATGGGCCATACCTAATAG 329
|||||
66 gHlsValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
|||||
328 GCATGTTGATCACTGAGGAGGAGAGGAGGTTGGAGATTCTAGGGC 279
|||||
83 lnglyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyr 99
|||||
278 AGGTGATGAAGCAAGCAAGAGTCTAAGGAGCAACAGATGAGCGCTTAC 229
|||||
100 ArgArgPheArgThrProGluProAsnHisThrAspPheCysLeuIle 116
|||||
228 ACCGTTTTCGAACCCCTGACCTGACATCATTTACGACTTTGCCCTCAT 179
|||||
116 ePro 117
|||||
178 ACCT 175

```

seq_name: gb_est2.BI289546

```

seq_documentation_block: 543 bp mRNA EST 19-JUL-2001
LOCUS BI289546
DEFINITION UI-R-DK0-cff-c-12-0-UI.81 UI-R-DK0 Rattus norvegicus cDNA clone
UI-R-DK0-cff-c-12-0-UI 3', mRNA sequence.
ACCESSION BI289546
VERSION BI289546
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 643)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized rat kidney pool library cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 1-51,
>AT_richLow complexity
Seq primer: M13 Forward
POLYA=yes. Location/Qualifiers
1. .643
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DK0-cff-c-12-0-UI"
/clone_lib="UI-R-DK0"
/clone_tag="ADUIT"
/lab_host="DHL08 (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DK0

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

```

```

FEATURES
source

```

Library is a subtracted library derived from a mixture of five individually tagged normalized rat libraries: brain-nRBP (20%), heart-nRBP (20%), kidney-nRBP (20%), aorta-nRBP (20%), and placenta-nRBP (20%). Each original library was constructed from a mixture of equal amounts of RNA from seven different developmental time-points: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, and adult day 200. (Exception: the aorta pool does not contain embryonic day 17 RNA and the placenta pool contains only the three embryonic stages). Each library was normalized individually according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the DK0 subtracted library, plasmid DNA from each of the five individually tagged normalized libraries was mixed in the proportions specified above and electroporated into competent bacteria for production of single-stranded circular DNA representing the pool of libraries. Single-stranded circular DNA representing these five normalized libraries was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a set of about 1,000 arrayed clones from each of the five non-normalized libraries of brain (CT0s), heart (CS0s), kidney (CU0s), aorta (CW0s), and placenta (CX0s). The resulting pool of approximately 5,000 clones represented about 33.3% of the final driver population. A set of about 2,000 arrayed clones from each of the five normalized libraries of brain (CT0), heart (CS0), kidney (CU0), aorta (CW0), and placenta (CX0). The resulting pool of about 10,000 clones represented about 66.6% of the final driver population.

```

TAG-LIB-UI-R-DK0
TAG-TISSUE-rat kidney pool
TAG-SEQ=CAAGACTGTC"
BASE COUNT 160 a 149 c 113 g 221 t
ORIGIN

```

```

alignment_scores:
Quality: 511.50 Length: 118
Ratio: 4.608 Gaps: 1
Percent Similarity: 94.068 Percent Identity: 86.441

```

```

alignment_block:
US-09-327-750D-35 x BI289546/rev ..
Align seg 1/1 to reverse of: BI289546 from: 1 to: 643

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
601 ATGGCGTCCAAAGTCAACAAAGTCATCTGGATCTCCTGTGGAGAAAG 552
|||||
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluLup 33
|||||
551 CAAGAAAAACAAAAAGGTGGAG.GCCTCCAAACAAAGTCAAGAGAAT 503
|||||
33 roHisHisLeuGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
502 CCCACATCTGGAAGAGTTGAACAAAGAGCCTGGGGCAATGTCAGG 453
|||||
50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
|||||
452 AGGAAGTCAAGCGACTTGTGCTAACTTCTATGGGCCATACCTAATAG 403
|||||
66 gHlsValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
|||||
402 GCATGTTGATCACTGAGGAGGAGAGGTTGGAGATTCTAGGGC 353
|||||
83 lnglyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyr 99
|||||
352 AGGTGATGAAGCAAGCAAGTCTAAGGAGCAACAGATGAGCGCTTAC 303

```

100 ArgArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeu11 116
 302 AGCGCTCCGACCCCTGAACCTGACATCATTTAGCACTTTTGCCTCAT 253
 116 ePro 117
 252 ACCT 249

seq_name: gb_est1:AA117429

seq_documentation_block:
 LOCUS AA117429 312 bp mRNA EST 15-NOV-1996
 DEFINITION mn2le12.r1 Beddington mouse embryonic region Mus musculus cDNA
 clone IMAGE:538606 5', similar to SW:HG74_HUMAN O00994_OVARIAN
 GRANULOSA CELL 13.0 KD PROTEIN HGR74. ;, mRNA sequence.
 ACCESSION AA117429
 VERSION AA117429.1 GI:1672442
 SOURCE EST.
 ORGANISM house mouse.
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 312)
 Author: Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:325542
 Seq primer: -40ml3 ET
 High quality sequence stop: 287.
 Location/Qualifiers
 1. 312

FEATURES

source
 /organism="Mus musculus"
 /strain="C57BL/6 x DBA"
 /db_xref="taxon:10090"
 /clone="IMAGE:538606"
 /sex="pooled"
 /tissue_type="embryo"
 /dev_stage="7.5dpc"
 /lab_host="DH12S"
 /note="Organ: whole embryo; Vector: pCMV-SPORT; Site: 1:
 SalI; Site: 2: NotI; Cloned unidirectionally. Primer:
 Oligo dt. Gastrulating embryos were collected at 7.5dpc
 from C57BL/6 x DBA matings, excluding embryos that had
 developed head folds and all extraembryonic tissues.
 Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
 Referenced in Development 121, 2479-2489 (1995)."
 BASE COUNT 110 a 56 c 92 g 54 t
 ORIGIN

alignment_scores:

Quality: 503.50 Length: 100
 Ratio: 5.086
 Percent Similarity: 99.000 Percent Identity: 98.000
 Gaps: 1

alignment_block:

US-09-327-750D-35 x AA117429

Align seg 1/1 to: AA117429 from: 1 to: 312

1 MetaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
 12 ATGGCATCCAAATTTAAACAAGTCATCTGGATCTCACTGTGGAGAAGA 61
 16 pLysLysAspLysArgGlyLysAlaSerLysGlnSerGluGluGluP 33
 62 CAAAAAGACAAAAAAGGTGGAGGCTCCAAACAAGTGAAGAAGAAC 111
 33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
 112 CCACCATCTGGAAGAGGTGAAACAAGAGAGCTGGGGGAAATGTCCGA 161
 50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
 162 AGGAAGATCAGCGACTTGTGCCTAACTTCTCTGGGCCATACCTAATG 211
 66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
 212 GCATGTGTATCCCAATGAAGGGGAGAGGATGTTGGAGATTTGTAGTGC 261
 83 InGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyr 99
 262 AGGGAACAGAACTCAAGAGAAAGACTACGAGAGCAGAGGTGAGGCCTTAC 311
 seq_name: gb_est2:W35893

seq_documentation_block:

LOCUS W35893 472 bp mRNA EST 14-MAY-1996
 DEFINITION mc53g09.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
 clone IMAGE:352288 5', similar to SW:HG74_HUMAN O00994_OVARIAN
 GRANULOSA CELL 13.0 KD PROTEIN HGR74. [1]; mRNA sequence.
 ACCESSION W35893
 VERSION W35893.1 GI:1317733
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 472)
 Author: Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:224088
 Seq primer: ETPrimer
 High quality sequence stop: 441.
 Location/Qualifiers
 1. 472

FEATURES

source
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:352288"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5]."
 BASE COUNT 110 a 56 c 92 g 54 t
 ORIGIN

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 465)
 Maria M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Willson, R. and Waterston, R.

The WashU-HMMI Mouse EST Project
 Unpublished (1996)
 Contact: Maria M/Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:229324

Seq primer: ETPRimer
 High quality sequence stop: 328.

FEATURES

source
 1. 465
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:367892"
 /clone_lib="Soares mouse embryo NBMEL3.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGCTACCAATCGAGTGGAGCGCGCGGAATTTTTTTTTTTTTTTTTTTT 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru KO, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 147 a 90 c 156 g 72 t

alignment_scores:
 Quality: 469.50 Length: 99
 Ratio: 4.840 Gaps: 1
 Percent Similarity: 97.980 Percent Identity: 96.970

alignment_block:

US-09-327-750D-35 x W54487

Align seg 1/1 to: W54487 from: 1 to: 465

1 MetaLaserLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
 168 ATGGCATCCAAATTTAAACAAGTCATACCTGATCTCCTGAGGAAAG 217
 16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGlu 33
 218 CAAAGAGACAAAAGAGTGGGAGGCGCTCCAAACAAGATGAAGAAGAC 267
 33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
 268 CCCACCATCTGGAAGAGTTGAAACAACAAGCCTGGGGAAATGTCCGA 317
 50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
 196 CAAAAAGACAAAAGAGTGGGAGGCGCTCCAAACAAGTGAAGAAGAC 245

318 AGAAGTC.AGGGACTTGTGCCTAACCTTCTCTGGGCCATACCAATAG 366
 66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
 367 GCATGTTGATCGCAATGAAGGGGAGAGATGTTGGGAGATTTGTAGTC 416
 83 lnglyThrGluValLysArgLysThrThrGluGlnGlnValArgPro 98
 417 AGGACACAGACTCAAGAGAAGACTACGGAGCAGCAGGTGAGGCCT 463

seq_name: gb_est1:AL118340

seq_documentation_block:

LOCUS AL118340 422 bp mRNA EST 23-SEP-1999
 DEFINITION v9112826 Beddington mouse dissected endoderm Mus musculus CDNA
 clone 528_13N20 5', mRNA sequence.

ACCESSION AL118340 GI:5920179

VERSION AL118340.1

KEYWORDS EST

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 422)
 AUTHORS Harrison, S.M., Dunwoodie, S.L., Arkell, R.M., Lehrach, H. and Beddington, R.S.

TITLE Isolation of novel tissue-specific genes from CDNA libraries representing the individual tissue constituents of the gastrulating mouse embryo
 JOURNAL Development 121 (8), 2479-2489 (1995)
 MEDLINE 95401865
 COMMENT Contact: Ruiz, P., Lehrach, H. and Avner, P.
 ESC Mouse Transcript Mapping Consortium
 Genoscope CNS
 2, rue Gaston Cremieux, CP 5706, 91057 Evry CEDEX, France
 Email: pavner@pasteur.fr
 Clone available from Ressourcenzentrum, Heubnerweg 6, D-14059 Berlin, Germany. Web site http://www.rzpd.de
 Seq primer: CCGTCCGGAATTCCTCCGGT.

FEATURES

source

1. 422
 /organism="Mus musculus"
 /strain="C57B16 x DBA"
 /db_xref="taxon:10090"
 /clone="528_13N20"
 /clone_lib="Beddington mouse dissected endoderm"
 /tissue_type="dissected endoderm"
 /dev_stage="7.5 dpc"
 /note="Vector: pSport1; Site 1: NotI; Site 2: Salt; Cloned unidirectionally. - High quality sequence only submitted. -

BASE COUNT 128 a 86 c 143 g 65 t
 ORIGIN

alignment_scores:
 Quality: 463.50 Length: 92
 Ratio: 5.093 Gaps: 1
 Percent Similarity: 98.913 Percent Identity: 98.913

alignment_block:

US-09-327-750D-35 x AL118340

Align seg 1/1 to: AL118340 from: 1 to: 422

1 MetaLaserLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
 146 ATGGCATCCAAATTTAAACAAGTCATACCTGATCTCCTGAGGAAAG 195
 16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGlu 33
 196 CAAAAAGACAAAAGAGTGGGAGGCGCTCCAAACAAGTGAAGAAGAC 245

6: 791-806, 1996.
TAG_LIB=UI-R-BJO
TAG_TISSUE=atrium at 16.5 dpc
TAG_SEQ=GATTC

BASE COUNT 146 a 120 c 94 g 180 t
ORIGIN

alignment_scores:
Quality: 459.00 Length: 97
Ratio: 5.044 Gaps: 0
Percent Similarity: 93.814 Percent Identity: 86.598

alignment_block:
US-09-327-750D-35 x AW251884/rev ..

Align seg 1/1 to reverse of: AW251884 from: 1 to: 540

21 ArgGlyGlyLysAlaSerLysGlnSerGluGluGluProHisLeuG1 37
:::|||||
538 AAAGTGGGAGGCTCCAAACAAAGTGAAGAGGATCCACCATCTGGA 489
37 uGluValGluAsnLysLysProGlyGlyAsnValArgLysValArgA 54
|||||
488 AGAGGTTGAAACAAAGAGGCTGGGGCAATGTCAGGAGGAAGTCAGGC 439
54 tGLeuValProAsnPhelLeuTipAlaIleProAsnArgHisValAspArg 70
|||||
438 GACTTGTGCTTAAGTCTTATGGCCATACCTAATAGGCATGTTGATCAC 389
71 AsnGluGlyGlyGluAspValGlyArgPheValValGlnGlyThrGluVa 87
:::|||||
388 AGTGAAGGGGAGAGGAGGTGGGAGATTCGTAGGCGAGGTGATGGAAGC 339
87 lLysArgLysThrThrGluGlnGlnValArgProTyrArgArgPheArgT 104
:|||||
338 CAAGAGAAGTCTAAGGAGCAACAGATGAGGCCTTACACGCGTTTCCGAA 289
104 hrProGluProAspAsnHisTyrAspPheCysLeuIlePro 117
|||||
288 CCCCTGAACCTGACAATCATTACGACCTTTTGCCTCATACCT 248

33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
246 CCCACCATCTGGAAGAGGTGAAACAAAGAGCCCTGGGGAAATGTCGGA 295
50 ArgLysValArgArgLeuValProAsnPhelLeuTipAlaIleProAsnAr 66
|||||
296 AGGAAAGTCAGGCGACTTGTGCTACTTCTCTGGCCATACCAATAG 345
66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValArg 83
|||||
346 GCATGTTTCATCGCAATGAGGGGAGAGGATGTTGGGAGATTGTCGTGC 395
83 lGlyThrGluValLysArgLysThr 91
|||||
396 AGGGAACAGAGTCAAGAGAAAGACT 421

seq_name: gb_est1:AW251884

seq_documentation_block:
LOCUS AW251884 540 bp mRNA EST 17-DEC-1999
DEFINITION UI-R-BJO-adn-b-06-0-UI-s1 UI-R-BJO Rattus norvegicus cDNA clone
UI-R-BJO-adn-b-06-0-UI 3', mRNA sequence.

ACCESSION AW251884
VERSION AW251884.1 GI:6595475
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 540)
Bonaldo M.F., Lennon G. and Soares M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized atrium at 16.5 dpc library cDNA Library Preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=yes.

FEATURES
source
1. .540
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJO-adn-b-06-0-UI"
/clone_lib="UI-R-BJO"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-BJO
library is a subtracted library derived from the UI-R-AAL,
UI-R-AB1, UI-R-AC1, UI-R-AE1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present between the Not I site
and the oligo-dT track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research

Tue Mar 12 09:01:46 2002

OM of: US-09-327-750D-35 to: N_Geneseq_1101:* out_format : pfs
Date: Mar 11, 2002 3:42 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=framet.p2n.model -DEV=xlh
-Q/cgn2_1/USP70.spool/US09327750/runat_11032002.101154.20362/app_query.fasta_1.1472
-DB=N_Geneseq_1101_QFMT-fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -YGAPOP=10.000 -YGAPEXT=0.500
-YGAPOP=6.000 -YGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09327750 -CGNL_1_330 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT_THREADS=1

Search information block:
Query: US-09-327-750D-35
Query length: 117
Database: N_Geneseq_1101:*
Database sequences: 930621
Database length: 428662619
Search time (sec): 355.560000

score	list:	Strd	Orig	ZScore	Escore	Len	Documentation
/SID2/gcgdata/geneseq/NA2001.DAT:AAH03517	+	355.50	767.81	1.1e-34	865		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	355.50	764.27	1.7e-34	1229		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	355.50	763.22	2.0e-34	1364		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	355.50	763.22	2.0e-34	1364		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	175.00	367.98	2.0e-12	891		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	174.00	365.47	2.0e-12	917		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	174.00	365.47	2.0e-12	917		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	153.00	321.71	7.6e-10	700		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	147.00	312.92	1.5e-09	485		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	141.00	297.91	1.5e-08	532		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	99.50	202.37	0.0034	767		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	99.50	201.19	0.0039	862		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	99.50	200.78	0.0042	898		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	98.50	159.28	0.8526	44453		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	94.00	189.87	0.0169	792		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	92.00	192.29	0.0124	401		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	89.00	189.41	0.0179	276		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	89.00	189.41	0.0179	276		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	89.00	189.41	0.0179	276		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	87.50	181.99	0.0453	415		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	87.50	181.99	0.0453	415		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	87.50	181.99	0.0453	415		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	87.50	177.29	0.0847	662		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	87.50	176.75	0.0907	698		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	84.50	174.68	0.1184	858		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	84.50	170.20	0.2101	692		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	82.00	157.98	0.101	1344		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	81.50	150.50	2.63	2533		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	80.50	118.25	164.41	50000		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	80.50	137.50	14.29	6755		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	79.50	149.29	3.07	1839		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	77.50	141.89	7.93	2471		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	76.00	123.33	75.44	14141		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	76.00	133.90	22.10	3929		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	76.00	129.63	38.21	6005		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	75.50	144.44	5.72	1236		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	75.50	129.01	41.37	5721		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	75.50	127.91	47.63	6381		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	75.50	143.96	6.08	1161		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	74.50	141.40	8.45	1342		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	74.50	141.39	8.46	1343		
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	74.50	127.88	47.82	5137		

/SID2/gcgdata/geneseq/NA1999.DAT:AA201022 - 74.50 103.74 1.1e+03 5652
/SID2/gcgdata/geneseq/NA2001.DAT:AA121493 + 74.00 132.63 2.00 394
/SID2/gcgdata/geneseq/NA2001.DAT:AA146784 + 74.00 132.63 2.00 394
/SID2/gcgdata/geneseq/NA2001.DAT:AA107190 + 74.00 132.63 2.00 394
/SID2/gcgdata/geneseq/NA1995.DAT:AA086237 + 74.00 128.11 46.45 4500

seq_name: /SID2/gcgdata/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:
ID AAH03517 standard; cDNA; 865 BP.
XX
AC AAH03517;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:352.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2007.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 03-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI; 2001-318749/34.
Primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
and/or diagnosis of the abnormality of the proteins encoded by the
full-length cDNAs -
Claim 1; SEQ ID 352; 2537pp + CD ROM; English.
The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dT primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the
oligonucleotide comprises at least 15 nucleotides and the combination of
the 5'-end sequence/3'-end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow obtaining of the full-length
cDNAs easily without any special methods. AAH03516 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
AAH5893 represent human amino acid sequences; and AAH13629 to AAH13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.
Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;